OM of: US-09-832-658-7 to: A_Geneseq_032802:* out_format = pir

Date: Oct 9, 2002 3:14 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL frame=_n2p.model -DEV=x11
-Q /cgn2_1/ptodata/[redacted]2658/app_query_09102002_083937_21417/app_query_fasta_1.144
-DB A_Geneseq_032802 -QFMT=fastan -SUFFIX=n2p.rag -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP 10.000 -XGAPEXT=0.500
-FGAPOP=6.000 FGAPEXT 7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP 6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS frame1to6.gdi -LIST=45 -DB=ALIGN 200 -THR_SCORE=pct
-THR_MAX 100 -THR_MIN=0 -ALIGN 15 -MODE LOCAL -OUTFMT=pir
-[redacted]M=ext HEAPSIZE=500 MINGEN 3 -MAXGEN=2000000000
-USER US09832658_3SEN1_1_60 -NCPU 6 -ICPU=4 -LONGLOG
-DEV_TIMEOUT 120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-832-658-7
Query length: 87
Database: A_Geneseq_032802:*
Database sequences: 747574
Database length: 111074796
Search time (sec): 34.620000

Score_list:
Sequence Strd Orig Pscore EScore Ion 1 Documentation

/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84920 + 124.00 459.20 4.0e-12
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY70865 + 124.00 459.20 4.0e-12
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84956 + 124.00 459.20 4.0e-12
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY70872 + 124.00 459.20 4.0e-12
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1982.DAT:AAP26026 + 91.00 269.24 1.5e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAB00069 + 89.00 267.71 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAB00172 + 89.00 267.71 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAB00177 + 89.00 267.71 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2001.DAT:AAB00184 + 89.00 267.71 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2001.DAT:AAB00185 + 89.00 267.71 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1981.DAT:AAP14056 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1984.DAT:AAP50224 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50024 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50026 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50027 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50028 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50029 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50030 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50031 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50032 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1985.DAT:AAP50033 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1986.DAT:AAP50028 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1986.DAT:AAP51071 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1986.DAT:AAP51072 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1986.DAT:AAP51073 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1986.DAT:AAP51074 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1987.DAT:AAP52028 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1988.DAT:AAP51012 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1988.DAT:AAP80047 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1988.DAT:AAP82007 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1988.DAT:AAW81734 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA1988.DAT:AAW49767 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84944 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84945 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84954 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84955 + 89.00 264.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84956 + 87.00 254.08 2.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84957 + 89.00 254.08 3.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84958 + 89.00 254.08 3.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84959 + 89.00 254.08 3.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84960 + 89.00 254.08 3.2e-06
/SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84961 + 89.00 254.08 3.2e-06

seq_name: /SIDS1/gcqdata/geneseq/geneseqp-embl/AA2000.DAT:AAY84920

seq_documentation_block:
ID AAY84920 standard; Protein; 183 AA.
XX
AC AAY84920;
XX
DT 21-AUG-2000 (first entry)
XX
DE Amino acid sequence of a His-tagged interferon beta 1a protein.
XX
KW Interferon beta 1a; IFN-beta1a; polymer; polyalkylene glycol; tumour;
KW cancer; autoimmune condition; fibrosis; lupus; multiple sclerosis;
KW viral disease; angiogenic disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..4
FT /label= 9
FT /note= "His tag"
FT Peptide 10..12
FT /note= "spacer"
FT Protein 18..183
FT /note= "natural interferon-betala"
FT
XX
PN WO200023114-A2.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-US024291.
XX
PR 15-OCT-1998; 98US-0104572.
PR 25-FEB-1999; 99US-0121961.
XX
PA (BIOT-) BIOGEN INC.
XX
PI Pepinsky R, Rutkel L, Sheppard M, Wiley A, Beckman P;
XX
DR WPI; 2000-356641/29.
DR N-PSDB; AAA15062.
XX
PT New glycosylated interferon-beta 1a coupled to a non naturally
PT occurring polymer containing a polyalkylene glycol useful for treating
PT e.g. tumours, autoimmune disorders, viral infections and angiogenic
PT diseases.
XX
PS Example 1; Fig 10; 74pp; English.
XX
CC The present sequence represents a glycosylated interferon beta 1a
CC (IFN-betala) protein. The protein is used to produce a composition
CC comprising a glycosylated IFN-beta coupled to a non naturally-occurring
CC polymer contained a polyalkylene oxide moiety. The composition has enhanced
CC activity, measured by an antiviral assay, as compared to unglycosylated
CC and/or IFN-beta. The polymer IFN-betala conjugate has the ability to
CC stay in the circulation for longer periods of time. It is useful in
CC stability in solution, reduced immunogenicity, protease resistance and
CC IFN-betala like properties. It is also useful in the treatment of viral
CC and increased thermal stability. The polymer based separates are useful
CC for treating tumours and cancer, as well as autoimmune conditions such
CC as fibrosis, lupus and multiple sclerosis. These may also be used in the
CC treatment of viral diseases, as well as in the treatment of angiogenic
CC diseases. IFN-betala is useful as an agent for the treatment, reduction
CC or alleviation of a disease state, physiological condition, symptoms or
CC pathological factors. It is therapeutically effective. The IFN-betala
CC polymer conjugates may also be used in prophylaxis or treatment of any

CC Condition or disease state for which the IFN beta-1a constituent is
CC efficacious, in biological systems or specimens.
XX
SQ Sequence 183 AA;

alignment_scores:
Quality: 124.00 Length: 28
Ratio: 5.167 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 82.143

alignment_block:
US-09-832-658-7 x AAY84920 ..

Align seq 1/1 to: AAY84920 from: 1 to: 183

3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGC 52
12 GlyAspAspAspAspLysMetSerTyrAsnLeuLeuGlyPheLeuGlnAr 28

53 TTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
28 qSerSerAsnPheGlnCysGlnLysLeuLeuTrp 59

Seq_name: /cgi-2/data_access_seq_pro_seq-2.1:AA2000:DAI:AAY70866

seq_documentation_block:
ID AAY70866 standard; Protein; 183 AA.
XX
AC AAY70866;
XX
DT 01-JUL-2000 (first entry)
XX
DE Histidine tagged human interferon beta 1a fusion protein.
XX
KW Human; interferon beta 1a; IFN beta 1a; immunoglobulin; fusion protein;
KW multiple sclerosis; antiinflammatory; immunosuppressive;
KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;
KW multiple sclerosis; inflammatory disease; autoimmune disease; cancer;
KW hepatitis; viral infection; neovascularisation; histidine tag; His tag.
XX
OS Homo sapiens
OS Synthetic
XX
FH Key Location/Qualifiers
FT Region 1..3
FT /note= "N-terminal residues remaining after cleavage of
FT vascular cell adhesion molecule-1 signal sequence"
FT Region 4..9
FT /label= His-tag
FT Region 10..12
FT /label= Spacer
FT Region 13..17
FT /label= Enterokinase linker
FT Protein 18..183
FT /label= natural human IFN-beta-1a
XX
PN WO2000024417-A2.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24201.
XX
PR 16-OCT-1998; 98US-0104491.
PR 16-FEB-1999; 99US-0120237.
XX
PA (BIOG-) BIOGEN INC.
XX
PI Worthy A, Pankert S, Brinkmann M, Hochman P.
XX
DR WPI; 2000-390654/29.
DR N-PSDB; AAI00165.
XX
PT Fusion proteins comprising interferon-beta 1a useful for inhibiting
PT angiogenesis -
XX
PS Example 1; Fig 1; 82pp; English.
XX
CC The patent discloses fusion proteins comprising glycosylated
CC interferon-beta (IFN-beta) especially IFN-beta-1a, linker groups and
CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein. The
CC fusion protein is useful for inhibiting angiogenesis in a patient.
CC It may also be used to treat multiple sclerosis, fibrosis, inflammatory
CC and autoimmune diseases, cancers, hepatitis and viral infection
CC characterised by neovascularisation. The present sequence is
CC a fusion protein that comprises histidine tag, an enterokinase
CC linker and human IFN-beta-1a. The His tagged IFN-beta fusion construct
CC was used as a template for mutagenesis in the construction of various
CC mutant His tagged-IFN-beta expression plasmids.
XX
SQ Sequence 183 AA;

alignment_scores:
Quality: 124.00 Length: 28
Ratio: 5.167 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 82.143

alignment_block:
US-09-832-658-7 x AAY70866 ..

Align seq 1/1 to: AAY70866 from: 1 to: 183

3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGC 52
12 GlyAspAspAspAspLysMetSerTyrAsnLeuLeuGlyPheLeuGlnAr 28

53 TTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
28 qSerSerAsnPheGlnCysGlnLysLeuLeuTrp 59

Seq_name: /cgi-2/data_access_seq_pro_seq-2.1:AA2000:DAI:AAY84950

seq_documentation_block:
ID AAY84950 standard; Protein; 166 AA.
XX
AC AAY84950;
XX
DT 21-AUG-2000 (first entry)
XX
DE Alanine mutant of human interferon beta 1a protein.
XX
KW Interferon beta 1a; IFN beta 1a; polymer; polyalkylene glycol; tumour;
KW cancer; autoimmune condition; fibrosis; lupus; multiple sclerosis;
KW viral disease; angiogenic disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "wild type Ser replaced with Ala"
FT Misc-difference 4
FT /note= "wild type Asn replaced with Ala"
FT Misc-difference 5
FT /note= "wild type Leu replaced with Ala"
FT Misc-difference 8
FT /note= "wild type Phe replaced with Ala"
FT Misc-difference 11
FT /note= "wild type Arg replaced with Ala"
XX
PN WO200024114 A2.
XX
PD 27 APR 2000.
XX
PF 15-OCT-1999; 99WO-US24201.

XX
PR 16-OCT-1998; 98US-0104572.
PR 16-FEB-1999; 99US-0120161.
XX
PA (BIOJ) BIOGEN INC.
XX
PI Pepinsky B, Runkel L, Brickelmaier M, Whitty A, Hochman P;
XX
DR WPI; 2000-339534/29.
XX
PT New glycosylated interferon-beta-1a coupled to a non-naturally
PT occurring polymer containing a polyalkylene glycol useful for treating
PT e.g. tumors, autoimmune disorders, viral infections and angiogenic
PT diseases -
XX
PS Example 1; Page -; 73pp; English.
XX
CC AAY84950-64 represent alanine mutants of human interferon beta 1a
CC (IFN-betala) protein. The protein is used to produce a composition
CC comprising a glycosylated IFN-beta coupled to a non-naturally-occurring
CC polymer containing a polyalkylene glycol. The composition has enhanced
CC activity, measured by an antiviral assay, as compared to physiologically
CC active IFN beta. The polymer-IFN-betala conjugate has the ability to
CC stay in the vasculature for longer periods of time, has increased
CC stability in solution, reduced immunogenicity, protection of the
CC modified IFN-beta-1a from proteolytic digestion and subsequent abolition
CC activity, and increased thermal stability. The polymer-based conjugates
CC are useful for treating tumors and cancer, as well as autoimmune
CC conditions such as fibrosis, lupus and multiple sclerosis. These may also
CC be used in the treatment of viral diseases, as well as in the treatment
CC of angiogenic diseases. IFN-betala is useful as an agent for the
CC treatment, remission or attenuation of a disease state, physiological
CC condition, symptoms or etiological factors, or for their evaluation or
CC diagnosis. The IFN-betala polymer conjugates may also be used for
CC prophylaxis or treatment of any condition or disease state for which the
CC IFN-betala constituent is efficacious, in biological systems or
CC specimens. note: the present sequence does not appear in the
CC specification; it was created using information provided.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 114.00 Length: 22
Ratio: 5.182 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment_block:
US-09-832-658-7 x AAY84950 ..

Alignment Seq 1/1 to: AAY84950 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
1 MetAlaTyrAlaAlaLeuGlyAlaLeuGlnAlaSerSerAsnPheGlnCy 17

71 PCAGAAGCTCCTGTGC 86
17 sGlnLysLeuLeuTrp 22

Seq_name: /SIS2/petit/Seqdata/1/AA2000.PAT.AAY70872

Seq_documentation_block:
ID AAY70872 standard; Protein; 166 AA.
XX
AC AAY70872;
XX
DT 31-JUL-2000 (first entry)
XX
DE Human interferon-beta alanine substituted mutant A1.
XX
KW human; interferon-beta; IFN-beta; immunoglobulin; fusion protein; mutant;
KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive;
KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;
KW multiple sclerosis; inflammatory disease; autoimmune disease; cancer;
KW hepatitis; viral infection; neovascularisation; IFN beta 1a; helix A.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..25
FT /label= Helix_A
FT Misc-difference 2
FT /note= "Wild type Ser is substituted by Ala"
FT Misc-difference 4
FT /note= "Wild type Asn is substituted by Ala"
FT Misc-difference 5
FT /note= "Wild type Leu is substituted by Ala"
FT Misc-difference 8
FT /note= "Wild type Phe is substituted by Ala"
FT Misc-difference 11
FT /note= "Wild type Arg is substituted by Ala"
FT Domain 26..53
FT /label= AB_loop
FT Domain 54..73
FT /label= Helix_B
FT Domain 74..100
FT /label= Helix_C
FT Domain 101..119
FT /label= CD_loop
FT Domain 120..134
FT /label= Helix_D
FT Domain 135..140
FT /label= DE_loop
FT Domain 141..166
FT /label= Helix_E
XX
PN WO200023472-A2.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24200.
XX
PR 16-OCT-1998; 98US-0104491.
PR 16-FEB-1999; 99US-0120237.
XX
PA (BIOJ) BIOGEN INC.
XX
PI Whitty A, Runkel L, Brickelmaier M, Hochman P;
XX
DR WPI; 2000-339654/29.
XX
PT Fusion proteins comprising interferon-beta-1a useful for inhibiting
PT angiogenesis -
XX
PS Example 1; Page ; 82pp; English.
XX
CC The patent discloses fusion proteins comprising glycosylated
CC interferon-beta (IFN-beta) especially IFN-beta-1a, linker groups and
CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein. The
CC fusion protein is useful for inhibiting angiogenesis in a patient.
CC It may also be used to treat multiple sclerosis, fibrosis, inflammatory
CC and autoimmune diseases, cancers, hepatitis and viral infection
CC characterised by neovascularisation. The present sequence is
CC a human interferon-beta alanine substituted mutant A1 generated by
CC mutating residues in the helix A of wild type sequence.
CC The Ala/Ser substituted mutants of IFN-beta having substitutions in
CC helices A, B, C, D or E, or loops AB, CD or DE were assessed for
CC receptor binding and functional activities e.g. antiviral and
CC antiproliferative activities. The A1 mutant shows antiviral and
CC antiproliferative activities which are disproportionately low with
CC respect to receptor binding, compared to wild-type IFN-beta-1a. The
CC mutant can be used to produce IFN-beta fusion proteins.
CC Note: The present sequence is not shown in the specification but is
CC derived from wild type human IFN-beta sequence found in page 46

```
(AAY70872).
XX
Sequence  166 AA;


Alignment scores:
            Quality:  214.000              Length:      22
              Ratio:    5.182              Gaps:         0
 Percent Similarity:  100.000   Percent Identity:  100.000

Matching blocks:
US-09-832-658-7 x AAY70872  ..

 Align seq 1/1  to: AAY70872  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTGGAGCCTACAAGCTTCTAGCAATTTCAGTG 70
          ||||||||||||||||||||||||||||||||||||||||||||||
        1 MetAlaTyrAlaAlaLeuGlyAlaLeuGlnAlaSerSerAsnPheGln'y 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

Seq_name: /tmp/biosci/tmp.pep.seq_gene.seq.pep.rel:AAzzzz...AT:AAU00069

Seq_documentation_block:
ID   AAU00069 standard; Protein; 166 AA.
XX
AC   AAU00069;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human interferon beta mutant N4K.
XX
KW   Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW   viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW   Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW   ulcerative colitis; immunomodulation; mutant; mutein; N4K.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Wild type Asn replaced with Lys"
XX
PN   WO200115736-A2.
XX
PD   08-MAR-2001.
XX
PF   25-AUG-2000; 2000WO-DK00471.
XX
PR   27-AUG-1999;   99DK-0001197.
PR   21-OCT-1999;   99US-0160782.
PR   26-NOV-1999;   99DK-0001691.
PR   07-FEB-2000; 2000DK-0000194.
PR   07-MAR-2000; 2000DK-0000363.
PR   14-APR-2000; 2000DK-0000642.
XX
PA   (MAXY-) MAXYGEN APS.
XX
PI   Pedersen AH,  Schambye HT,  Andersen KV,  Bornaes C,  Rasmussen HB;
XX
DR   WPI; 2001-218488/22.
XX
PT   A conjugate exhibiting interferon beta activity useful for treating
PT   multiple sclerosis comprises a non-polypeptide group covalently
PT   attached to an interferon beta polypeptide
XX
PS   Claim 11; Page - ; 108pp; English.
XX
CC   The sequence represents Human interferon beta mutant N4K.
CC   Conjugates of the invention exhibiting interferon beta activity
CC   comprise at least one first non polypeptide group covalently attached
CC   to an interferon beta polypeptide, the amino acid sequence of which
CC   differs from wild-type human interferon beta in at least one introduced
CC   and at least one removed amino acid residue comprising an attachment
CC   group for the first non-polypeptide group. The invention also concerns
CC   reducing the immunogenicity and/or increasing functional in vivo
CC   half-life and/or serum half life of an interferon beta polypeptide
CC   comprising introducing an amino acid residue constituting an attachment
CC   group for a first non-polypeptide group into a position exposed at the
CC   surface of the protein that does not contain such a group and removing
```

```
(AAY70872).
XX
Sequence  166 AA;


Alignment scores:
            Quality:  214.000              Length:      22
              Ratio:    5.182              Gaps:         0
 Percent Similarity:  100.000   Percent Identity:  100.000

Matching blocks:
US-09-832-658-7 x AAY70872  ..

 Align seq 1/1  to: AAY70872  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTGGAGCCTACAAGCTTCTAGCAATTTCAGTG 70
          ||||||||||||||||||||||||||||||||||||||||||||||
        1 MetAlaTyrAlaAlaLeuGlyAlaLeuGlnAlaSerSerAsnPheGln'y 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

Seq_name: /tmp/biosci/tmp.pep.seq_gene.seq.pep.rel:AAzzzz...AT:AAU00069

Seq_documentation_block:
ID   AAU00069 standard; Protein; 166 AA.
XX
AC   AAU00069;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human interferon beta mutant N4K.
XX
KW   Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW   viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW   Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW   ulcerative colitis; immunomodulation; mutant; mutein; N4K.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Wild type Asn replaced with Lys"
XX
PN   WO200115736-A2.
XX
PD   08-MAR-2001.
XX
PF   25-AUG-2000; 2000WO-DK00471.
XX
PR   27-AUG-1999;   99DK-0001197.
PR   21-OCT-1999;   99US-0160782.
PR   26-NOV-1999;   99DK-0001691.
PR   07-FEB-2000; 2000DK-0000194.
PR   07-MAR-2000; 2000DK-0000363.
PR   14-APR-2000; 2000DK-0000642.
XX
PA   (MAXY-) MAXYGEN APS.
XX
PI   Pedersen AH,  Schambye HT,  Andersen KV,  Bornaes C,  Rasmussen HB;
XX
DR   WPI; 2001-218488/22.
XX
PT   A conjugate exhibiting interferon beta activity useful for treating
PT   multiple sclerosis comprises a non-polypeptide group covalently
PT   attached to an interferon beta polypeptide
XX
PS   Claim 11; Page - ; 108pp; English.
XX
CC   The sequence represents Human interferon beta mutant N4K.
CC   Conjugates of the invention exhibiting interferon beta activity
CC   comprise at least one first non polypeptide group covalently attached
CC   to an interferon beta polypeptide, the amino acid sequence of which
CC   differs from wild-type human interferon beta in at least one introduced
CC   and at least one removed amino acid residue comprising an attachment
CC   group for the first non-polypeptide group. The invention also concerns
CC   reducing the immunogenicity and/or increasing functional in vivo
CC   half-life and/or serum half life of an interferon beta polypeptide
CC   comprising introducing an amino acid residue constituting an attachment
CC   group for a first non-polypeptide group into a position exposed at the
CC   surface of the protein that does not contain such a group and removing

Alignment scores:
            Quality:  214.000              Length:      22
              Ratio:    5.182              Gaps:         0
 Percent Similarity:  100.000   Percent Identity:  100.000

Matching blocks:
US-09-832-658-7 x AAY70872  ..

 Align seq 1/1  to: AAY70872  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTGGAGCCTACAAGCTTCTAGCAATTTCAGTG 70
          ||||||||||||||||||||||||||||||||||||||||||||||
        1 MetAlaTyrAlaAlaLeuGlyAlaLeuGlnAlaSerSerAsnPheGln'y 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

alignment_block:
 US-09-832-658-7 x AAY70872  .

 Align seq 1/1  to: AAY70872  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTGGAGCCTACAAGCTTCTAGCAATTTCAGTG 70
          ||||||||||||||||||||||||||||||||||||||||||||||
        1 MetAlaTyrAlaAlaLeuGlyAlaLeuGlnAlaSerSerAsnPheGln'y 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

Seq_name: /tmp/biosci/tmp.pep.seq_gene.seq.pep.rel:AAzzzz...AT:AAU00069

Seq_documentation_block:
ID   AAU00069 standard; Protein; 166 AA.
XX
AC   AAU00069;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human interferon beta mutant N4K.
XX
KW   Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW   viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW   Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW   ulcerative colitis; immunomodulation; mutant; mutein; N4K.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Wild type Asn replaced with Lys"
XX
PN   WO200115736-A2.
XX
PD   08-MAR-2001.
XX
PF   25-AUG-2000; 2000WO-DK00471.
XX
PR   27-AUG-1999;   99DK-0001197.
PR   21-OCT-1999;   99US-0160782.
PR   26-NOV-1999;   99DK-0001691.
PR   07-FEB-2000; 2000DK-0000194.
PR   07-MAR-2000; 2000DK-0000363.
PR   14-APR-2000; 2000DK-0000642.
XX
PA   (MAXY-) MAXYGEN APS.
XX
PI   Pedersen AH,  Schambye HT,  Andersen KV,  Bornaes C,  Rasmussen HB;
XX
DR   WPI; 2001-218488/22.
XX
PT   A conjugate exhibiting interferon beta activity useful for treating
PT   multiple sclerosis comprises a non-polypeptide group covalently
PT   attached to an interferon beta polypeptide
XX
PS   Claim 11; Page - ; 108pp; English.
XX
CC   The sequence represents Human interferon beta mutant N4K.
CC   Conjugates of the invention exhibiting interferon beta activity
CC   comprise at least one first non polypeptide group covalently attached
CC   to an interferon beta polypeptide, the amino acid sequence of which
CC   differs from wild-type human interferon beta in at least one introduced
CC   and at least one removed amino acid residue comprising an attachment
CC   group for the first non-polypeptide group. The invention also concerns
CC   reducing the immunogenicity and/or increasing functional in vivo
CC   half-life and/or serum half life of an interferon beta polypeptide
CC   comprising introducing an amino acid residue constituting an attachment
CC   group for a first non-polypeptide group into a position exposed at the
CC   surface of the protein that does not contain such a group and removing
```

CC an amino acid residue constituting an attachment group for a first
CC non-polypeptide group and subjecting the modified peptide to conjugation
CC with the non-polypeptide group. The conjugate and a cell culture
CC expressing the mutated polypeptides are useful in the treatment of
CC disease, especially multiple sclerosis, and for treating mammals having
CC circulating antibodies against interferon beta 1a or 1b. DNA encoding the
CC mutated proteins may be used for gene therapy. The DNA and proteins can
CC also be used to treat viral infections (e.g viral hepatitis), cancer
CC (e.g. breast cancer), inflammation, Crohn's disease, acute myeloid
CC leukaemia, Hodgkin's disease and ulcerative colitis and for
CC immunomodulation.
CC Note: The present sequence is not shown in the specification but is
CC derived from the human interferon beta sequence given in AAU00038.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 90.00 Length: 22
Ratio: 5.000 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x AAU00069 ..

Align seq 1/1 to: AAU00069 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
|||:::|||| |||||| |||||| ||||||||||||||||||
1 MetSerTyrLysLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 TCAGAAGCTCCTGTGG 86
||||||||||||||||
17 sGlnLysLeuLeuTrp 22

Seq_name: /SIDS1/gcgdata/geneseq/geneseqp-emb:/AA2001.DAT:AAU00069

Seq_documentation_block:
ID AAU00069 standard; Protein; 166 AA.
XX
AC AAU00069;
XX
DT 16-MAY-2001 (first entry)
XX
DE Human interferon beta mutant N4K/S27K/D39K/Q49K/E85K/A89K
XX
KW Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW ulcerative colitis; immunomodulation; mutant; mutein;
KW N4K/S27K/D39K/Q49K/E85K/A89K
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 4
FT /note= "Wild type Asn replaced with Lys"
FT Misc-difference 27
FT /note= "Wild-type Arg replaced with Lys"
FT Misc-difference 39
FT /note= "Wild-type Asp replaced with Lys"
FT Misc-difference 49
FT /note= "Wild-type Gln replaced with Lys"
FT Misc-difference 85
FT /note= "Wild-type Glu replaced with Lys"
FT Misc-difference 89
FT /note= "Wild-type Ala replaced with Lys"
XX
PN WO200115736-A2.
XX
PD 08-MAR-2001.
XX
PF 25-AUG-2000; 2000WO-DK00471.
XX
PR 27-AUG-1999; 99DK-0001197.
PR 21-OCT-1999; 99US-0160782.
PR 26-NOV-1999; 99DK-0001691.
PR 07-FEB-2000; 2000DK-0000194.
PR 07-MAR-2000; 2000DK-0000363.
PR 14-APR-2000; 2000DK-0000642.
XX
PA (MAXY-) MAXYGEN APS.
XX
PI Pedersen AH, Schambye HT, Andersen KV, Bornaes C, Rasmussen PB;
XX
DR WPI; 2001-218488/22
XX
PT A conjugate exhibiting interferon beta activity useful for treating
PT multiple sclerosis comprises a non-polypeptide group covalently
PT attached to an interferon beta polypeptide -
XX
PS Disclosure; Page - ; 108pp; English.
XX
CC The sequence represents Human interferon beta mutant
CC N4K/S27K/D39K/Q49K/E85K/A89K. Conjugates of the invention
CC exhibiting interferon beta activity comprise at least one
CC first non polypeptide group covalently attached to an
CC interferon beta polypeptide, the amino acid sequence of which
CC differs from wild-type human interferon beta in at least one introduced
CC and at least one removed amino acid residue comprising an attachment
CC group for the first non-polypeptide group. The invention also concerns
CC reducing the immunogenicity and/or increasing functional in vivo
CC half-life and/or serum half-life of an interferon beta polypeptide
CC comprising introducing an amino acid residue constituting an attachment
CC group for a first non-polypeptide group into a position exposed at the
CC surface of the protein that does not contain such a group and removing
CC an amino acid residue constituting an attachment group for a first
CC non-polypeptide group and subjecting the modified peptide to conjugation
CC with the non-polypeptide group. The conjugate and a cell culture
CC expressing the mutated polypeptides are useful in the treatment of
CC disease, especially multiple sclerosis, and for treating mammals having
CC circulating antibodies against interferon beta 1a or 1b. DNA encoding the
CC mutated proteins may be used for gene therapy. The DNA and proteins can
CC also be used to treat viral infections (e.g viral hepatitis), cancer
CC (e.g. breast cancer), inflammation, Crohn's disease, acute myeloid
CC leukaemia, Hodgkin's disease and ulcerative colitis and for
CC immunomodulation.
CC Note: The present sequence is not shown in the specification but is
CC derived from the human interferon beta sequence given in AAU00038.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 90.00 Length: 22
Ratio: 5.000 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x AAU00176 ..

Align seq 1/1 to: AAU00176 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
|||:::|||| |||||| |||||| |||||||||||||||||||
1 MetSerTyrLysLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 TCAGAAGCTCCTGTGG 86
||||||||||||||||
17 sGlnLysLeuLeuTrp 22

Seq_name: /SIDS1/gcgdata/geneseq/geneseqp-emb:/AA2001.DAT:AAU00177

seq_documentation_block:
ID AAU00177 standard; Protein; 166 AA.
XX

```
AC   AAU00177;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human interferon beta mutant N4E,K7E,K11R,Q49K,E85K,A89K,K123R
XX
KW   Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW   viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW   Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW   ulcerative colitis; immunomodulation; mutant; mutein;
KW   N4E,K7E,K11R,Q49K,E85K,A89K,K123R.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Wild-type Asn replaced with Lys"
FT   Misc-difference 27
FT                   /note= "Wild-type Arg replaced with Lys"
FT   Misc-difference 39
FT                   /note= "Wild-type Asp replaced with Lys"
FT   Misc-difference 49
FT                   /note= "Wild-type Gln replaced with Lys"
FT   Misc-difference 85
FT                   /note= "Wild-type Glu replaced with Lys"
FT   Misc-difference 89
FT                   /note= "Wild-type Ala replaced with Lys"
XX
PN   WO200115736-A2.
XX
PD   08 MAR-2001.
XX
PF   25-AUG-2000; 2000WO-DK00471.
XX
PR   27-AUG-1999;   99DK-0001197.
PR   21-OCT-1999;   99US-0160782.
XX
CC   circulating antibodies against interferon beta 1a or 1b. DNA encoding the
CC   mutated proteins may be used for gene therapy. The DNA and proteins can
CC   also be used to treat viral infections (e.g. viral hepatitis), cancer
CC   (e.g. breast cancer), inflammation, Crohn's disease, acute myeloid
CC   leukaemia, Hodgkin's disease and ulcerative colitis and for
CC   immunomodulation.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from the human interferon beta sequence given in AAU00008.
XX
SQ   Sequence   166 AA;


alignment_scores:
                Quality:   90.00              Length:      22
                  Ratio:   5.000                Gaps:       0
 Percent Similarity:  81.818   Percent Identity: 77.273

alignment_block:
 US-09-832-658-7 x AAU00177  ..

 Align_seq 1/1  to: AAU00177  from: 1  to: 166

       21 ATGGCTTACGCCCTCTTCGCAATCGGTATAACCTTCTTAGCAATTTTCAGTG 70
          |||:::|||      |||||  ||||||  |||||||||||||||||
        1 MetSerTyrLysLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

seq_name:  ... /genetics/bio_seq/bio_seq.cgi?cmd=AAZwOlLtAr:AAu00184

seq_documentation_block:
ID   AAU00184 standard; Protein; 166 AA.
XX
AC   AAU00184;
XX
DT   16-MAY-2001  (first entry)
XX
DE   Human interferon beta mutant N4E,K7E,K11R,Q49K,E85K,A89K,K123R.
XX
KW   Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW   viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW   Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW   ulcerative colitis; immunomodulation; mutant; mutein;
KW   N4E,K7E,K11R,Q49K,E85K,A89K,K123R.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /note= "Wild-type Asn replaced with Lys"
FT   Misc-difference 27
FT                   /note= "Wild-type Arg replaced with Lys"
FT   Misc-difference 39
FT                   /note= "Wild type Asp replaced with Lys"
FT   Misc-difference 49
FT                   /note= "Wild-type Gln replaced with Lys"
FT   Misc-difference 85
FT                   /note= "Wild type Glu replaced with Lys"
FT   Misc-difference 89
FT                   /note= "Wild type Ala replaced with Lys"
FT   Misc-difference 123
FT                   /note= "Wild type Lys replaced with Arg"
XX
PN   WO200115736-A2.
XX
PD   08 MAR-2001.
XX
PF   25-AUG-2000; 2000WO-DK00471.
XX
PR   27-AUG-1999;   99DK-0001197.
PR   21-OCT-1999;   99US-0160782.
```

PR 26-NOV-1999; 99DK-0001691.
PR 07-FEB-2000; 2000DK-0000194.
PR 07-MAR-2000; 2000DK-0000363.
PR 14-APR-2000; 2000DK-0000642.
XX
PA (MAXY-) MAXYGEN APS.
XX
PI Pedersen AH, Schambye HT, Andersen KV, Bornaes C, Rasmussen PB;
XX
DR WPI; 2001-218488/22.
XX
PT A conjugate exhibiting interferon beta activity useful for treating
PT multiple sclerosis comprises a non-polypeptide group covalently
PT attached to an interferon beta polypeptide -
XX
PS Disclosure; Page - ; 108pp; English.
XX
CC The sequence represents Human interferon beta mutant
CC N4K,E27K,D96K,K97R,S98R,K99R,Q49R,T85K,A89K,K123R. Conjugates of the invention
CC exhibiting interferon beta activity comprise at least one
CC first non-polypeptide group covalently attached to an interferon
CC beta polypeptide, the amino acid sequence of which differs from
CC wild-type human interferon beta in at least one introduced and at
CC least one removed amino acid residue comprising an attachment
CC group for the first non-polypeptide group. The invention also concerns
CC reducing the immunogenicity and/or increasing functional in vivo
CC half-life and/or serum half-life of an interferon beta polypeptide
CC comprising introducing an amino acid residue constituting an attachment
CC group for a first non-polypeptide group into a position exposed at the
CC surface of the protein that does not contain such a group and removing
CC an amino acid residue constituting an attachment group for a first
CC non-polypeptide group and subjecting the modified peptide to conjugation
CC with the non-polypeptide group. The conjugate and a cell culture
CC expressing the mutated polypeptides are useful in the treatment of
CC disease, especially multiple sclerosis, and for treating mammals having
CC circulating antibodies against interferon beta 1a or 1b. DNA encoding the
CC mutated proteins may be used for gene therapy. The DNA and proteins can
CC also be used to treat viral infections (e.g. viral hepatitis), cancer
CC (e.g. breast cancer), inflammation, Crohn's disease, acute myeloid
CC leukaemia, Hodgkin's disease, and ulcerative colitis and for
CC immunomodulation.
CC Note: The present sequence is not shown in the specification but is
CC derived from the human interferon beta sequence given in AAU00038.
XX
SQ Sequence 166 AA;

Alignment_scores:
Quality: 90.00 Length: 22
Ratio: 5.000 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x AAU00184 ..

Align seq 1/1 to: AAU00184 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
| ||||:||| ||||:| |||||| |||||||||||||||||
1 MetSerTyrLysLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 TCACAAGCTCCTGTGC 86
|||||||||||||||||
17 sGlnLysLeuLeuVal 22

seq_name: /cgi21/sq-pdb/seq-fasta/home of embl/AA2001_DAT:AAU00185

seq_documentation_block:
ID AAU00185 standard; Protein; 166 AA.
XX
AC AAU00185;
XX
DT 16-MAY-2001 (first entry)

XX
DE Human interferon beta mutant N4K,E27K,D96K,K97R,S98R,K99R,Q49R,T85K,A89K,K123R.
XX
KW Human; interferon beta; antibody; multiple sclerosis; gene therapy;
KW viral infection; viral hepatitis; cancer; breast cancer; inflammation;
KW Crohn's disease; acute myeloid leukaemia; Hodgkin's disease;
KW ulcerative colitis; immunomodulation; mutant; mutein;
KW N4K/E27K/D96K/K97R/S98R/K99R/Q49R/T85K/A89K/K123R.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 4
FT /note= "Wild-type Asn replaced with Lys"
FT Misc-difference 27
FT /note= "Wild-type Arg replaced with Lys"
FT Misc difference 39
FT /note= "Wild-type Asp replaced with Lys"
FT Misc-difference 45
FT /note= "Wild-type Lys replaced with Arg"
FT Misc-difference 49
FT /note= "Wild-type Gln replaced with Lys"
FT Misc-difference 85
FT /note= "Wild-type Glu replaced with Lys"
FT Misc-difference 89
FT /note= "Wild type Ala replaced with Lys"
FT Misc difference 123
FT /note= "Wild-type Lys replaced with Arg"
XX
PN WO200115736-A2.
XX
PD 08-MAR-2001.
XX
PF 25-AUG-2000; 2000WO-DK00471.
XX
PR 27-AUG-1999; 99DK-0001197.
PR 21-OCT-1999; 99US-0160782.
PR 26-NOV-1999; 99DK-0001691.
PR 07-FEB-2000; 2000DK-0000194.
PR 07-MAR-2000; 2000DK-0000363.
PR 14-APR-2000; 2000DK-0000642.
XX
PA (MAXY-) MAXYGEN APS.
XX
PI Pedersen AH, Schambye HT, Andersen KV, Bornaes C, Rasmussen PB;
XX
DR WPI; 2001-218488/22.
XX
PT A conjugate exhibiting interferon beta activity useful for treating
PT multiple sclerosis comprises a non-polypeptide group covalently
PT attached to an interferon beta polypeptide -
XX
PS Disclosure; Page - ; 108pp; English.
XX
CC The sequence represents Human interferon beta mutant
CC N4K,E27K,D96K,K97R,S98R,K99R,Q49R,T85K,A89K,K123R. Conjugates of the invention
CC exhibiting interferon beta activity comprise at least one
CC first non-polypeptide group covalently attached to an interferon
CC beta polypeptide, the amino acid sequence of which differs from
CC wild-type human interferon beta in at least one introduced and at
CC least one removed amino acid residue comprising an attachment
CC group for the first non-polypeptide group. The invention also concerns
CC reducing the immunogenicity and/or increasing functional in vivo
CC half-life and/or serum half-life of an interferon beta polypeptide
CC comprising introducing an amino acid residue constituting an attachment
CC group for a first non-polypeptide group into a position exposed at the
CC surface of the protein that does not contain such a group and removing
CC an amino acid residue constituting an attachment group for a first
CC non-polypeptide group and subjecting the modified peptide to conjugation
CC with the non-polypeptide group. The conjugate and a cell culture
CC expressing the mutated polypeptides are useful in the treatment of
CC disease, especially multiple sclerosis, and for treating mammals having
CC circulating antibodies against interferon beta 1a or 1b. DNA encoding the

```
CC   mutated proteins may be used for gene therapy. The DNA and proteins can
CC   also be used to treat viral infections (e.g viral hepatitis), cancer
CC   (e.g. breast cancer), inflammation, Crohn's disease, acute myeloid
CC   leukaemia, Hodgkin's disease and ulcerative colitis and for
CC   immunomodulation.
CC   Note: The present sequence is not shown in the specification but is
CC   derived from the human interferon beta sequence given in AAU00038.
XX
SQ   Sequence   166 AA;


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:       0
 Percent Similarity:  81.818   Percent Identity: 77.273

alignment_block:
 US-09-832-658-7 x AAU00185   ..

 Align seq 1/1  to: AAU00185  from: 1  to: 166

      21 APGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
         |||::|||       ||||||   ||||||   ||||||||||||||||
       1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

      71 TCAGAAGCTCCTGTGG 86
         ||||||||||||||||
      17 sGlnLysLeuLeuTrp 22

Seq name: ...

Seq documentation block:
ID   AAP30219 Standard; Protein; 166 AA.
XX
AC   AAP30219;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -beta-1 encoded by plasmid
DE   pDM101/trp/beta-1.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
PN   WO8302461-A.
XX
PD   21 JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF, Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; AAN30152.
XX
PT   Multiclass hybrid interferon polypeptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 5; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it. Pref. the
CC   AA sequence consists of alpha and beta interferons. Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN alpha 1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN beta 1) (see AAN30155, AAP30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN beta 1) (see
CC   AAN30160, AAP30227)   Alternatively IF1 is the amino terminal end of a
CC   to the 5' terminal of human fibroblast IFN.  The entire sequence of
CC   mature human fibroblast interferon is encoded by AAN10049 (see also
CC   AAP10051).  Another recombinant IFN which was recovered differed from
CC   AAN10049 in that the codon for the 49th amino acid had a "silent"
CC   base change (TAC to TAT), indicating the existence of genic
CC   polymorphism (see AAP10052/N10051).
XX
SQ   Sequence   31 AA;


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:       0
 Percent Similarity:  81.818   Percent Identity: 77.273

alignment_block:
 US-09-832-658-7 x AAP10052   ..

 Align seq 1/1  to: AAP10052  from: 1  to: 31

      21 APGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
         |||::|||       ||||||   ||||||   ||||||||||||||||
       1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

      71 TCAGAAGCTCCTGTGG 86
         ||||||||||||||||
      17 sGlnLysLeuLeuTrp 22

Seq name: ...

Seq documentation block:
ID   AAP30219 Standard; Protein; 166 AA.
XX
AC   AAP30219;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -beta-1 encoded by plasmid
DE   pDM101/trp/beta-1.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
PN   WO8302461-A.
XX
PD   21 JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF, Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; AAN30152.
XX
PT   Multiclass hybrid interferon polypeptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 5; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it. Pref. the
CC   AA sequence consists of alpha and beta interferons. Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN alpha 1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN beta 1) (see AAN30155, AAP30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN beta 1) (see
CC   AAN30160, AAP30227)   Alternatively IF1 is the amino terminal end of a
```

CC beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC 1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC resp.) (see AAN30156, AAN30223). In the examples plasmids pGW5 and
CC pDM101/trp/beta 1 and p alpha-61A were used (see AAN30151, AAN30152,
CC AAN30157). HintI was used to digest the DNA sequences in the region
CC of significant handicaps (see AAN30153, AAN30154, AAN30158, AAN30159),
CC and the restriction fragments were ligated to form hybrid DNA.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 89.00 Length: 22
Ratio: 4.944 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x AAP30219 ..

Align seq 1/1 to: AAP30219 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 TCAGAAGCTCCTGTGG 86
17 sGlnLysLeuLeuTrp 22

seq_name: /SIDS1/gcgdata/geneseq/geneseqp/embl/AA1985.DAT:AAP30219

seq_documentation_block:
ID AAP30219 standard; Protein; 166 AA.
XX
AC AAP30219;
XX
DT 04-SEP-1991 (first entry)
XX
DE Sequence of new modified human beta interferon polypeptides
DE IFNX 418.
XX
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative.
XX
OS Homo sapiens.
XX
PN EP163993-A.
XX
PD 11-DEC-1985.
XX
PF 17-MAY-1985; 85EP-0105750.
XX
PR 17-MAY-1984; 84GB-0012564.
XX
PA (SEAR) SEARLE G D & CO.
XX
PI Bell LD, Boseley PG, Porter AG;
XX
DR WPI; 1985-311944/50.
DR N-PSDB; AAN50025.
XX
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
XX
PS Claim 16; Chart 2c; page 34; 71pp; English.
XX
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targetting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 89.00 Length: 22
Ratio: 4.944 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x AAP50024 ..

Align seq 1/1 to: AAP50024 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 TCAGAAGCTCCTGTGG 86
17 sGlnLysLeuLeuTrp 22

seq_name: /SIDS1/gcgdata/geneseq/geneseqp/embl/AA1985.DAT:AAP50026

seq_documentation_block:
ID AAP50026 standard; Protein; 166 AA.
XX
AC AAP50026;
XX
DT 04-SEP-1991 (first entry)
XX
DE Sequence of new modified human beta interferon polypeptides
DE IFNX 444.
XX
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative.
XX
OS Homo sapiens.
XX
PN EP163993-A.
XX
PD 11-DEC-1985.
XX
PF 17-MAY-1985; 85EP-0105750.
XX
PR 17-MAY-1984; 84GB-0012564.
XX
PA (SEAR) SEARLE G D & CO.
XX
PI Bell LD, Boseley PG, Porter AG;
XX
DR WPI; 1985-311944/50.
DR N-PSDB; AAN50027.
XX
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
XX
PS Claim 18; Chart 2e; page 36; 71pp; English.
XX
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targetting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
XX
SQ Sequence 166 AA;

alignment_scores:
Quality: 89.00 Length: 22
Ratio: 4.944 Gaps: 0

Percent Similarity: 81.818 Percent Identity: 77.273

Alignment block:
US-09-832-658-7 x AAB50026 ..

Align Seq 1/1 to: AAB50026 from: 1 to: 166

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAAIIIICAGIG 70
||||::|||| |||||| ||:||| ||||||||||||||||||
1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

71 ICAGAAGCICCIGIGG 86
||||||||||||||||
17 sGlnLysLeuLeuTrp 22

Result from Align_n_Block:
Query: AAB50027.A standard; Protein; 166 AA.
XX
AC AAB50027;
XX
04 SEP 1998 (first entry)
XX
DE Sequence of new modified human beta interferon polypeptides
DE IFNX 445.
XX
KW Antiviral; cell growth regulator; immune system regulator;
KW multiple sclerosis.
XX
OS Homo sapiens.
XX
PN EP836606-A
XX
PD 22 APR 1998.
XX
PF 27 MAY 1997; 97EP-923529.
XX
PR 17 MAY 1996; 96GB-0012504.
XX
PA (GENE) GENELABS TECH INC.
XX
PI Bell LD, Bosselay PG, Porter AG;
XX
DR WPI; 1998-211441/19.
DR N-PSDB; AAN50028.
XX
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti proliferative and immune regulating actions
XX
PS Claim 19; Page 47; 71pp; English.
XX
CC Compared with interferon beta prepd. by recombinant methods, the
CC IFNs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targetting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
XX
SQ Sequence 166 AA;

Alignment scores:
Quality: 80.80 Length: 22
Ratio: 1.944 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

Alignment block:
US-09-832-658-7 x AAB50027 ..

Align Seq 1/1 to: AAB50027 from: 1 to: 166

```
OM 01: US-09-832-658-7 to: PIR_71:*  -out_format = pfs

Date: Oct 9, 2002  3:15 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=framer_n2p.model -DEV=x11h
 -Q=/cgn2_16/prod/pub/[illegible]/app_query.fasta_1.144
 -DB PIR_71 -QFMT fastan -SUFFIX=n2p.rpr -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP 4.500  QGAPEXT=0.250  -XGAPOP=10.000 -XGAPEXT=0.500
 -PGAPOP 6.000  PGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000  DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -[illegible]KM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -[illegible]R=US09832658_2CGN1_1_52 -NCPU=6 -ICPU=3 -LONCLOG
 -[illegible]_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLIPXY -WAIT -THREADS=1

Search Information block:
Query: US-09-832-658-7
Query length: 87
Database: PIR_71:*
Database sequences: 284138
Database length: 96089334
Search time (sec): 23.450000

score_list:
Sequence                    Strd Orig    EScore    EScore Len  ! Documentation  ..
pir1:IVHUB1                   +   89.00   250.10  4.1e-06   187  ! interferon beta 1 precursor [va
pir2:I11189                   +   53.00   145.53  4.05      459  ! NADH dehydrogenase (ubiquinone)
pir2:AB3568                   -   50.50   121.78  10.58    1025  ! acriflavin resistance protein
pir2:T20327                   +   50.00   127.32  12.76     417  ! hypothetical protein D1586.6
pir2:JN0809                   -   49.50   122.37  15.49     649  ! drebrin E (clone qBbH13) - huma
pir2:T14763                   +   49.50   122.34  15.49     651  ! hypothetical protein DKFZp434D0
pir2:S60588                   +   49.50   121.70  15.49     707  ! drebrin A - rat
pir2:D71528                   +   49.00   134.84  18.62     109  ! probable pts phosphocarrier pro
pir2:A97337                   -   49.00   128.89  18.68     233  ! ABC-type transporter, ATPase co
pir2:E98198                   +   49.00   126.41  18.70     300  ! epxD protein (D36097) [imported
pir2:AE3088                   -   49.00   126.41  18.70     300  ! polysaccharide export protein b
pir2:AD2066                   +   49.00   126.41  18.70     320  ! cobalamin biosynthetic protein
pir2:AB2255                   +   49.00   119.89  18.76     736  ! po ypkosphate kinase [imported]
[illegible]:T11168            +   48.50   124.30  22.66     346  ! NADH dehydrogenase (ubiquinone)
[illegible]:JC7217            -   48.00   125.77  27.40     237  ! paternally imprinted protein PE
pir2:A71165                   -   48.00   125.29  27.41     252  ! hypothetical protein PH0518 - P
pir2:B36470                   -   48.00   122.50  27.45     360  ! Wnt-2 protein - mouse
pir2:A41672                   -   48.00   115.99  27.54     827  ! glycerol-3-phosphate acyltransf
pir2:B90614                   +   47.50   121.31  33.24     346  ! NADH dehydrogenase chain 2 ND2
pir2:B90616                   -   47.50   121.31  33.24     346  ! NADH dehydrogenase chain 2 ND2
pir2:B90624                   -   47.50   121.31  33.24     346  ! NADH dehydrogenase chain 2 ND2
pir2:A43776                   -   47.50   116.92  33.31     607  ! drebrin E2 - chicken
pir2:S64408                   -   47.00   120.51  40.24     317  ! hypothetical protein MJ0871 - M
pir2:T42252                   +   47.00   118.29  40.29     421  ! polypeptide N-acetylgalactosami
pir2:S48416                   -   47.00   114.40  40.36     692  ! probable membrane protein YPL01
pir2:T42691                   -   47.00   113.35  40.38     791  ! hypothetical protein DKFZp434D2
pir1:A43100                   -   47.00   107.32  40.51    1708  ! ataxia telangiectasia-associat
pir2:D71072                   -   46.50   112.19  48.90     758  ! hypothetical protein PH1271 - P
pir2:T42207                   -   46.50   100.57  49.19    3343  ! breast cancer susceptibility p
pir2:G70627                   -   46.00   122.44  58.88     169  ! hypothetical protein Rv0268c -
pir2:AC0536                   -   46.00   122.01  58.89     178  ! probable lipoprotein [imported]
pir2:A90067                   +   46.00   119.90  58.96     234  ! hypothetical protein SA2294 [im
pir2:T04269                   -   46.00   118.32  59.02     286  ! hypothetical protein F20B18.240
pir2:F70766                   -   46.00   117.59  59.02     314  ! hypothetical protein Rv2093 - M
pir2:C72216                   -   46.00   117.23  59.04     329  ! endoglucanase - Thermotoga mari
pir2:S20723                   -   46.00   114.76  59.11     451  ! carboxypeptidase (EC 3.4.17.-)
pir2:B75469                   -   46.00   114.42  59.12     477  ! multidrug-efflux transporter -
pir2:T20752                   -   46.00   111.79  59.20     659  ! hypothetical protein F11A5.8 -
pir2:I12624                   -   46.00   111.40  59.21     702  ! NADH dehydrogenase (ubiquinone)
pir2:E91049                   -   46.00   110.81  59.22     747  ! probable cytochrome C-type biog
pir2:F65026                   +   46.00   110.81  59.22     747  ! hypothetical protein b2503
pir2:A85894                   +   46.00   110.81  59.22     747  ! probable cytochrome C-type b
pir2:T22595                   +   46.00   103.99  59.43    1785  ! hypothetical protein F53B4.
pir2:T18314                   +   46.00   103.54  59.44    1892  ! hypothetical protein L7610.
pir2:T47054                   +   45.50   121.00  71.31     168  ! hypothetical protein [import

seq_name: pir1:IVHUB1

seq_documentation_block:
interferon beta-1 precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 08-Dec-2000
C;Accession: A93721; A93887; A91468; A93225; A93706; A93700; A93269; S04479; 156415;
R;Lawn, R.M.; Adelman, J.; Franke, A.E.; Houck, C.M.; Gross, M.; Najarian, R.; Goeddo
Nucleic Acids Res. 9, 1045-1052, 1981
A;Title: Human fibroblast interferon gene lacks introns.
A;Reference number: A93721; MUID:81198952
A;Accession: A93721
A;Molecule type: DNA
A;Residues: 1-187 <LAW>
A;Cross-references: GB:V00535; NID:g32639; PIDN:CAA23796.1; PID:g32640
R;Ohno, S.; Taniguchi, T.
Proc. Natl. Acad. Sci. U.S.A. 78, 5305-5309, 1981
A;Title: Structure of a chromosomal gene for human interferon beta.
A;Reference number: A93887
A;Accession: A93887
A;Molecule type: DNA
A;Residues: 1-187 <OHN>
R;Taniguchi, T.; Ohno, S.; Fujii-Kuriyama, Y.; Muramatsu, M.
Gene 10, 11-15, 1980
A;Title: The nucleotide sequence of human fibroblast interferon cDNA.
A;Reference number: A91468; MUID:81005095
A;Accession: A91468
A;Molecule type: mRNA
A;Residues: 1-187 <TAN>
A;Cross-references: GB:V00546; NID:g32735; PIDN:CAA23807.1; PID:g32736
R;Derynck, R.; Content, J.; DeClercq, E.; Volckaert, G.; Tavernier, J.; Devos, R.; Fi
Nature 285, 542-547, 1980
A;Title: Isolation and structure of a human fibroblast interferon gene.
A;Reference number: A93225; MUID:80254542
A;Accession: A93225
A;Molecule type: DNA
A;Residues: 1-187 <DER>
A;Cross-references: GB:V00546; NID:g32735; PIDN:CAA23807.1; PID:g32736
R;Houghton, M.; Eaton, M.A.W.; Stewart, A.G.; Smith, J.C.; Doel, S.M.; Catlin, G.H.;
Nucleic Acids Res. 8, 2885-2894, 1980
A;Title: The complete amino acid sequence of human fibroblast interferon as deduced u
A;Reference number: A93706; MUID:81053720
A;Accession: A93706
A;Molecule type: mRNA
A;Residues: 1-187 <HOU>
A;Cross-references: GB:J00218; GB:K00616; GB:M11029; NID:g184626; PID:g386802
R;Houghton, M.; Stewart, A.G.; Doel, S.M.; Emtage, J.S.; Eaton, M.A.W.; Smith, J.C.;
Nucleic Acids Res. 8, 1913-1931, 1980
A;Title: The amino-terminal sequence of human fibroblast interferon as deduced from t
A;Reference number: A93700; MUID:81053854
A;Accession: A93700
A;Molecule type: mRNA
A;Residues: 1-68 <HOU2>
R;Wetzel, R.
Nature 289, 606-607, 1981
A;Title: Assignment of the disulphide bonds of leukocyte interferon.
A;Reference number: A93244; MUID:81123083
A;Contents: annotation; disulfide bond
R;Shepard, H.M.; Leung, D.; Stebbing, N.; Goeddel, D.V.
Nature 294, 563-565, 1981
A;Title: A single amino acid change in IFN-beta 1 abolishes its antiviral activity.
A;Reference number: A93269; MUID:82080683
A;Accession: A93269
A;Molecule type: mRNA
A;Residues: 71-161,'Y',163-187 <SHE>
A;Experimental source: variant, clone pF526
A;Note: the loss of Cys-162 (and of the ability to form the essential disulfide bond)
```

R;[illegible], T.; Mizuno, Y.; Bosoi, K.; Ohkubo, K.; Sawada, R.; Kajitani, M.; Sakai, T.; Nar
Eur. J. Biochem. 181, 545-553, 1989
A;Title: Interferons and oncogenes: evolution of four different mammalian cell-derived recombinant human int
A;Reference number: S04479; MUID:89276336
A;Accession: S04479
A;Status: preliminary
A;Molecule type: protein
A;Residues: 22-207 <HIS>
R;Marx, L.L.; Schgal, P.B.
J. Interferon Res. 5, 521-526, 1985
A;Title: Is the relationship between human interferon alpha-1 and beta-1 genes.
A;Reference number: I96315; MUID:86035565
A;Accession: I96315
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-187 <HIS>
A;Cross-references: GB:M28622; NID:g184624; PIDN:AAA36010.1; PID:g184625
C;Genetics:
A;Gene: GDB:IFN-alpha; IFNB
A;Cross-references: GDB:120522; OMIM:147640
A;Map position: 9p21-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral
F;1-22/Domain: signal sequence #status predicted <SIG>
F;1-187/Product: interferon beta-1 #status experimental <MAT>
F;23-162/Disulfide bonds: #status predicted

alignment scores:
Quality: [illegible] Length: 22
Ratio: 4.444 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment block:
US-09-832-658-7 x [illegible]

Align seq 1/1 to: [illegible] from: 1 to: 187

21 A[illegible]AG 70
[illegible]
22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCys 88

71 [illegible] 86
[illegible]
89 sTrpLysGlyAla 99

seq_name: pir2:[illegible]

seq_documentation_block:
[illegible] dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Smithornis sharpei mitochondrion
C;Species: mitochondrion Smithornis sharpei
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
C;Accession: I11280
R;Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A;Title: Multiple independent origins of mitochondrial gene order in birds.
A;Reference number: Z17242
A;Accession: I11280
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-459 <MIN>
A;Cross-references: EMBL:AF090340; NID:g4894488; PID:g4894498; PIDN:AAD32523.1
C;Genetics:
A;Gene: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane associated complex; mitochondrion; NAD; oxidative phosphorylation;

alignment scores:
Quality: [illegible] Length: 17
Ratio: 1.417 Gaps: 0
Percent Similarity: [illegible] Percent Identity: 58.824

alignment block:
US-09-832-658-7 x FI1189 ..

Align seq 1/1 to: FI1189 from: 1 to: 459

7 ACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCT 56
[illegible]
36 ThrIleSerLeuGlnTrpLeuThrProThrLeuTyrProTyrLysAsnLe 52

57 A 57
|
52 u 52

seq_name: pir2:AH3568

seq_documentation_block:
acriflavin resistance protein F [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AH3568
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, D.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AB3252; PMID:11756688
A;Accession: AH3568
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1025 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53715.1; PID:g17984638; GSPDB:GN00150
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0473
A;Map position: II
C;Superfamily: hypothetical protein b2075

alignment scores:
Quality: 50.50 Length: 42
Ratio: 2.405 Gaps: 1
Percent Similarity: 65.625 Percent Identity: 37.500

alignment block:
US-09-832-658-7 x AH3568 ..

Align seq 1/1 to: AH3568 from: 1 to: 1025

1 CCGGACACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA 50
[illegible]
56 ProAspThrMetAlaSerSerValAlaThrProLeuLeuLysGlnPheGl 72

51 GCTTCTAGCA.....ATTTTCAGTGTCAGAAGCTCCTGTGGC 87
[illegible]
72 uThrIleAlaGlyIleAspThrIleSerAlaArgSerSerLeuGly 87

seq_name: pir2:T20327

seq_documentation_block:
hypothetical protein D1086.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20327
R;Smye, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19257
A;Accession: T20327
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-417 <WIL>
A;Cross-references: EMBL:Z81491; PIDN:CAB04823.1; GSPDB:GN00023; CESP:D1086.6
A;Experimental source: clone D1086
C;Genetics:

A;Gene: CESP:D10086.6
A;Map position: 5
A;Introns: 40/3, 171/3, 217/3, 332/3, 373/3, 777/3, 1072/3

aligment_scores:
Quality: 50.00 Length: 24
Ratio: 2.941 Gaps: 0
Percent Similarity: 70.833 Percent Identity: 45.833

aligment_block:
US-09-832-658-7 x I20327 ..

Align seq 1/1 to: I20327 from: 1 to: 417

6 GACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTC 55
48 AspAspValLysArgGlnCysAlaAlaAlaGlyGlyAsnTyrTyrAlaGl 59

56 TAGCAATTTTCAGTGTCAGAAG 77
59 yAlaAsnPheGlnCysSerLys 86

Seq_name: pir2:JN0809

seq_documentation_block:
drebrin E (clone qDbhl3) - human
C;Species: Homo sapiens (man)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C;Accession: JN0809
R;Ioda, M.; Shirao, T.; Minoshima, S.; Shimizu, N., Toya, S.; Uyemura, K.
Biochem. Biophys. Res. Commun. 196, 468-472, 1993
A;Title: Molecular cloning of cDNA encoding human drebrin E and chromosomal mapping of t
A;Reference number: JN0809; MUID:94030036
A;Accession: JN0809
A;Molecule type: mRNA
A;Residues: 1-649 <TOD>
A;Cross-references: DDBJ:D17530; NID:g498650; PIDN:BAA04480.1; PID:g498651
A;Experimental source: brain
C;Comment: This protein is an actin-binding protein.
C;Genetics:
A;Map position: 5
C;Keywords: actin binding

aligment_scores:
Quality: 49.50 Length: 29
Ratio: 2.152 Gaps: 1
Percent Similarity: 79.310 Percent Identity: 41.379

aligment_block:
US-09-832-658-7 x JN0809 ..

Align seq 1/1 to: JN0809 from: 1 to: 649

3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnG 53

51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
53 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 65

seq_name: pir2:T14763

seq_documentation_block:
hypothetical protein DKFZp434D064.1 - human
C;Species: Homo sapiens (man)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14763
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z18181
A;Accession: T14763
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-651 <WAM>
A;Cross-references: EMBL:AL110225
A;Experimental source: adult testis; clone DKFZp434D064
C;Genetics:
A;Note: DKFZp434D064.1

aligment_scores:
Quality: 49.50 Length: 29
Ratio: 2.152 Gaps: 1
Percent Similarity: 79.310 Percent Identity: 41.379

aligment_block:
US-09-832-658-7 x T14763 ..

Align seq 1/1 to: T14763 from: 1 to: 651

3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
39 GlySerAspAspLeuLysLeuAlaAlaSerGlyGlnGlyGlyLeuGlnGl 55

51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
55 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 67

seq_name: pir2:S60588

seq_documentation_block:
drebrin A - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C;Accession: S60588
R;Shirao, T.; Kojima, N.; Obata, K.
Neuroreport 3, 109-112, 1992
A;Title: Cloning of drebrin A and induction of neurite-like processes in drebrin tran
A;Reference number: S60588, MUID:92305233
A;Accession: S60588
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-707 <SHI>
A;Cross-references: EMBL:X59267; NID:g297820; PIDN:CAA41957.1; PID:g297821

aligment_scores:
Quality: 49.50 Length: 29
Ratio: 2.152 Gaps: 1
Percent Similarity: 79.310 Percent Identity: 41.379

aligment_block:
US-09-832-658-7 x S60588 ..

Align seq 1/1 to: S60588 from: 1 to: 707

3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 53

51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
53 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 65

seq_name: pir2:D71528

seq_documentation_block:
probable pts phosphocarrier protein hpr - Chlamydia trachomatis (serotype D, strain U
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: D71528
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998

A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809
A;Accession: D71528
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <ABN>
A;Cross-references: GB:AE001273; GB:AE001273; NID:g3329718; PIDN:AAC67932.1; PID:g3329
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pldB
C;Function:
Suggestively: phosphoinositidase-system phosphohistidine-containing protein homology
Pir2:D71528 major phosphoinositidase-system phosphohistidine-containing protein homology

alignment_scores:
Quality: 44.00 Length: 19
Ratio: 4.267 Gaps: 0
Percent Similarity: 78.947 Percent Identity: 47.368

alignment_block:
US-09-832-658-7 x D71528 ..

Align seq 1/1 to: D71528 from: 1 to: 109

61 CTCTCTTTCCAGCTCAAGCTTCTAGCAATTTTCAGTGTCAGAAGCTC 80
||||||:|||| :|||||| |||:|||||||||||||||
16 ProLeuMetGluAspGlyGluLeuSerAlaValPheThrIleArgAsnSe 32

81 CTGTCGT 87
: ||
32 rSerGly 34

seq_name: pir2:AG7337

seq_documentation_block:
ABC-type transporter, ATPase component CAC3560 [imported] - Clostridium acetobutylicum
Organism: Clostridium acetobutylicum
Update: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: AG7337
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Makarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Dubois, T.; Tatusov, R.L.; Koonin, E.V.; Wolf, Y.I.; Galperin, M.Y.; Fraser, C.M.;
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: AG6000; MUID:21599725; PMID:21599825
A;Accession: AG7337
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK81484.1; PID:g7026655; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3560

alignment_scores:
Quality: 44.00 Length: 14
Ratio: 4.308 Gaps: 0
Percent Similarity: 85.714 Percent Identity: 71.429

alignment_block:
US-09-832-658-7/rev x AG7337 ..

Align seq 1/1 to: AG7337 from: 1 to: 233

66 CAAAATTCTACAAGCTTCTAGCTCTCAACAGCGGCCTAAG 25
|||||| ||:::|||||||||:::||||||||||||
134 GluAsnPheLeuLysPheValGlyLeuGluGluArgArgLys 147

seq_name: pir2:E98198

seq_documentation_block:
pxbD protein (EC-like) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
Organism: Agrobacterium tumefaciens
Update: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C;Accession: E98198
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: E98198
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-300 <KUR>
A;Cross-references: GB:AL583870; GB:AAK87111.1; PID:g15158415; GSDB:GN00170
C;Genetics:
A;Gene: AGR_L_1067
A;Map position: linear chromosome

alignment_scores:
Quality: 49.00 Length: 20
Ratio: 2.882 Gaps: 0
Percent Similarity: 85.000 Percent Identity: 45.000

alignment_block:
US-09-832-658-7 x E98198 ..

Align seq 1/1 to: E98198 from: 1 to: 300

9 GATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAG 58
||| ||| :::|||:::|||||||||||||:::::::::::
163 AspGlyAspAlaAlaSerTyrThrAlaLeuGlyAlaPheLysSerAlaGl 179

59 CAATTTTCAG 68
::::||||||
179 yGluPheGln 182

seq_name: pir2:AE3088

seq_documentation_block:
polysaccharide export protein bexD [imported] - Agrobacterium tumefaciens (strain C58
Organism: Agrobacterium tumefaciens
Update: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AE3088
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, C.; Cillet, W.; Grant, C.; Creenhaer, P.; Katygarin, T.; Levy, R.; Li, M.; Mecl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AE3088
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-300 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45123.1; PID:g17427794; GSDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: bexD
A;Map position: linear chromosome

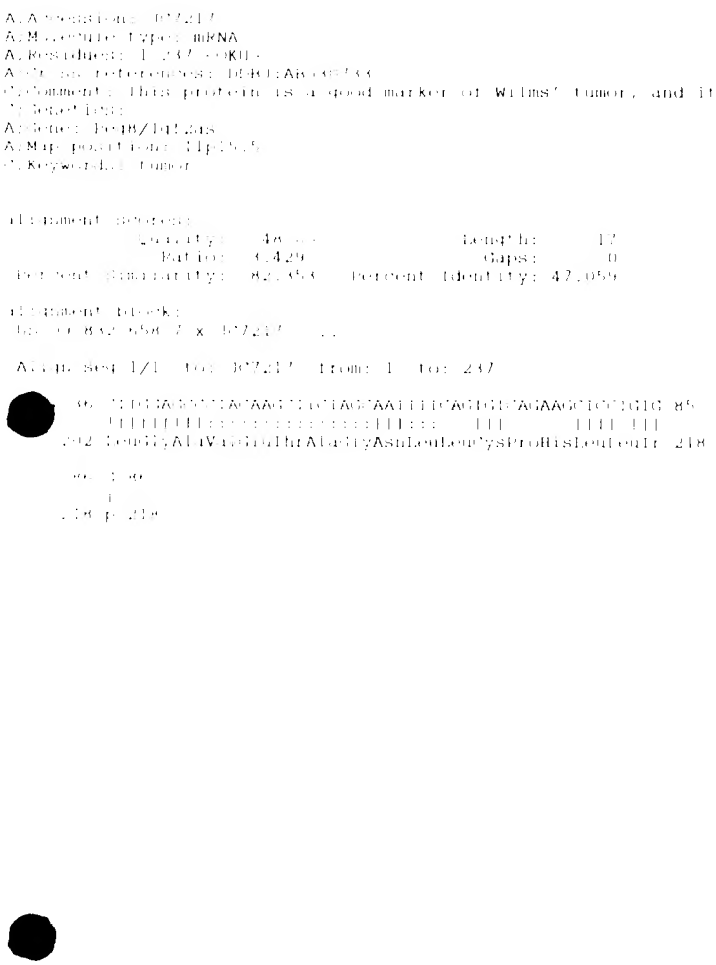
alignment_scores:
Quality: 49.00 Length: 20
Ratio: 2.882 Gaps: 0
Percent Similarity: 85.000 Percent Identity: 45.000

alignment_block:
US-09-832-658-7 x AE3088 ..

Align seq 1/1 to: AE3088 from: 1 to: 300

9 GATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAG 58
||| ||| :::|||:::|||||||||||||:::::::::::

```
     164 AspGlyAspAlaAlaSerTyrThrAlaLeuGlyAlaPheLysSerAlaGl 179

      59 CAATTTCAG 68
         :::::||||
     179 yGluPheGln 182

seq_name: pir2:AD2066

seq_documentation_block:
cobalamin biosynthetic protein [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AD2066
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB73781.1; PID:g17131173; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: cobD
C;Superfamily: cobalamin biosynthesis protein D


alignment scores:
            Quality:   49.00              Length:      26
              Ratio:   2.882               Gaps:        1
 Percent Similarity:   65.385   Percent Identity: 42.308

alignment_block:
 US-09-832-658-7 x AD2066    ..

 Align seq 1/1  to: AD2066  from: 1  to: 320

      13 AtGACAAGATGGCTTACGCCGCT.............TTCGAGCCCTACAA 50
         |||||:|||||||||||:::|||:::           ||||||
      77 MetThrArgTrpLeuAsnProValLeuAlaIleAlaLeuGluSerIleMe 93

      51 GCTTCTAGCAATTTTCAGTCTCACAAGC 78
         |||   :::   |||::::::::|||
      93 tLeuAlaSerCysPheAlaLeuLysSer 102

seq_name: pir2:AB2255

seq_documentation_block:
polyphosphate kinase [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AB2255
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2255
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-736 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB75292.1; PID:g17132726; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3593
C;Superfamily: Escherichia coli polyphosphate kinase


alignment scores:
            Quality:   49.00              Length:      14
              Ratio:   4.083               Gaps:        0
 Percent Similarity:   85.714   Percent Identity: 78.571

alignment_block:
 US-09-832-658-7 x AB2255    ..

 Align seq 1/1  to: AB2255  from: 1  to: 736

      28 ACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGT 69
         ||||||||||||||||      |||:::|||||||||||||||
      42 ThrProLeuLeuGluArgLeuLysPheLeuAlaIlePheSer 55

seq_name: pir2:T11168

seq_documentation_block:
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - greater rhea mitochondrion
C;Species: mitochondrion Rhea americana (greater rhea, common rhea)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C;Accession: T11168; T11417
R;Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A;Title: Multiple independent origins of mitochondrial gene order in birds.
A;Reference number: Z17242
A;Accession: T11168
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-346 <MIN>
A;Cross-references: EMBL:AF090339; NID:g4894475; PID:g4894477; PIDN:AAD32504.1
R;Harlid, A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z17270
A;Accession: T11417
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-44,'T',46-343,'T',345-346 <HAR>
A;Cross-references: EMBL:Y16884; PIDN:CAA76502.1
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylatio


alignment scores:
            Quality:   48.50              Length:      23
              Ratio:   2.853               Gaps:        1
 Percent Similarity:   73.913   Percent Identity: 47.826

alignment_block:
 US-09-832-658-7 x T11168    ..

 Align seq 1/1  to: T11168  from: 1  to: 346

       7 ACGATGATGACAAGATGGCTT...ACGCCGCTCTTGGAGCCCTACAAGCT 53
         ||||||||||||   |||      |||||:::|||:::         ::
     229 ThrMetMetThrSerTrpThrLysThrProIleLeuAsnAlaThrLeuMe 245

      54 TCTAGCAATTTTCACTCTC 72
         :|||||||:::::|||:::
     245 tLeuAlaLeuLeuSerLeu 251

seq_name: pir2:JC7217

seq_documentation_block:
paternally imprinted protein PEG8/IGF2AS [imported] - human
C;Species: Homo sapiens (man)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 18-Aug-2000
C;Accession: JC7217
R;Okutsu, T.; Kuroiwa, Y.; Kagitani, F.; Kai, M.; Aisaka, K.; Tsutsumi, O.; Kaneko, Y
J. Biochem. 127, 475-483, 2000
A;Title: Expression and imprinting status of human PEG8/IGF2AS, a paternally expresse
A;Reference number: JC7217
```



```
SM of: US-09-832-658-7 to: SwissProt_40:*   out_format : pfs

Date: Oct 9, 2002  3:16 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
  MODEL frame=_n2p.model -DEV=xlh
 -Q /cgn2_10/data/[illegible]/app_query.fasta_1.144
  DB SwissProt_40 -QFMT=fastan -SUFFIX=n2p.rsp -GAPOP 12.000
 -GAPEXT 4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP 10.000 -XGAPEXT=0.500
 -FGAPOP 6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
  DELOP 6.000 -DELEXT=7.000 -START 1 -MATRIX blosum62
 -TRANS human1@.edi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
  THR_MAX 100 -THR_MIN=0 -ALIGN 15 -MODE LOCAL -OUTFMT=pfs
 -[illegible]RM ext  HEAPSIZE=500  MINLEN 0  MAXLEN=2000000000
 -[illegible]ER US09832658_2CGN1_1_19 -NCPU=6 -ICPU=3 -LONGLOG
  DEV_TIMEOUT 120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-832-658-7
Query length: 87
Database: SwissProt_40:*
Database sequences: 105224
Database length: 38719550
Search time (sec): 14.990000

score_list:
Sequence                  Strd Orig    ZScore    EScore Len  ! Documentation  ..
SwissProt_40:INB_HUMAN +    89.00  246.12  2.8e-06   187  ! P01574 homo sapiens (human).
SwissProt_40:DREB_HUMAN +    49.50  121.10  7.34       649  ! Q16643 homo sapiens (human).
SwissProt_40:DREB_RAT +    49.50  120.44  7.33       707  ! Q07266 rattus norvegicus (rat)
SwissProt_40:PIHP_CHLTR +    49.00  133.47  8.95       109  ! O84341 chlamydia trachomatis
SwissProt_40:EYA4_CHICK +    48.50  131.33  10.79      119  ! Q9yh98 gallus gallus (chicke
SwissProt_40:WNT2_MOUSE -    48.00  121.29  12.92      360  ! P21552 mus musculus (mouse).
SwissProt_40:PLSB_MOUSE -    48.00  114.85  12.85      827  ! Q61586 mus musculus (mouse).
SwissProt_40:PLSB_HUMAN +    48.00  114.84  12.85      828  ! O94hc2 homo sapiens (human)
SwissProt_40:PLSB_RAT +    48.00  114.84  12.85      828  ! P97564 rattus norvegicus (rat)
SwissProt_40:CREP_CHICK +    47.50  115.23  15.52      652  ! P18302 gallus gallus (chicke
SwissProt_40:11IB_TRIVU +    47.00  120.63  18.82      269  ! Q9xs77 trichosurus vulpecul
SwissProt_40:7871_METJA +    47.00  119.35  18.80      317  ! Q58281 methanococcus jannasc
SwissProt_40:YF83_YEAST +    47.00  118.31  18.71      692  ! P43579 saccharomyces cerevis
SwissProt_40:AIM_HUMAN +    47.00  103.80  18.53     3056  ! Q13315 homo sapiens (human).
SwissProt_40:EYA4_FUGRU +    46.50  125.48  22.81      119  ! Q9yha1 fugu rubripes (japan
SwissProt_40:EYA4_MOUSE +    46.50  112.75  22.58      616  ! Q9z191 mus musculus (mouse).
SwissProt_40:EYA4_HUMAN +    46.50  112.46  22.57      639  ! O95677 homo sapiens (human).
SwissProt_40:YK83_MYCTU +    46.00  116.51  27.34      314  ! P18143 streptomyces griseus.
SwissProt_40:CRPS_STRGR +    46.00  113.70  27.28      451  ! P27172 escherichia coli. hyp
SwissProt_40:YEGF_ECOLI +    46.00  109.79  27.19      747  ! O43592 homo sapiens (human).
SwissProt_40:XPOT_HUMAN -    46.00  107.83  27.15      962  ! P97429 mus musculus (mouse)
SwissProt_40:BR02_MOUSE -    45.50   96.75  32.48     3329  ! O21515 phodopus sungorus (su
SwissProt_40:ND4M_PHOSU +    45.00  116.78  39.87      208  ! P27006 xenopus laevis (afri
SwissProt_40:ANXB_XENLA +    45.00  112.99  39.75      339  ! P49337 gallus gallus (chicke
SwissProt_40:WN14_CHICK -    45.00  112.72  39.74      351  ! P56705 homo sapiens (human).
SwissProt_40:WN14_HUMAN -    45.00  112.72  39.74      351  ! P22724 mus musculus (mouse).
SwissProt_40:WN14_MOUSE      45.00  112.72  39.74      351  ! Q9qxq5 rattus norvegicus (rat)
SwissProt_40:WNT4_RAT -    45.00  112.72  39.74      395  ! P31392 rattus norvegicus (rat)
SwissProt_40:ADMR_RAT +    45.00  111.81  39.71      404  ! [illegible]
SwissProt_40:TH11_BACHD +    45.00  111.63  39.71      707  ! P16788 human cytomegalovirus
SwissProt_40:G3VK_HCMVA +    45.00  107.30  39.56      728  ! Q55898 synechocystis sp. (str
SwissProt_40:PPK_SYNY3 +    45.00  107.07  39.56      743  ! P14923 homo sapiens (human).
SwissProt_40:PLAK_HUMAN +    45.00  106.91  39.55     1004  ! P28774 artemia santrancisca
SwissProt_40:ATIA_ARTSF +    45.00  104.58  39.48     1172  ! O64474 arabidopsis thaliana
SwissProt_40:AHM2_AKATH +    45.00  103.38  39.44     2516  ! Q24270 drosophila melanogas
SwissProt_40:CAD_DROME +    45.00   97.46  39.25      219  ! Q9ftz2 oryza sativa (rice). P
SwissProt_40:FHP_ORYSA +    44.50  114.92  48.07      277  ! P24200 escherichia coli  5-m
SwissProt_40:MCRA_ECOLI -    44.50  113.10  48.00      279  ! Q05559 rhizobium leguminosar
SwissProt_40:FIXA_RHILP +    44.50  113.04  48.00      923  ! P44781 haemophilus influe
SwissProt_40:HEPA_HAEIN +    44.50  103.77  47.63      967  ! P23852 escherichia coli.
SwissProt_40:HEPA_ECOLI +    44.50  103.41  47.62      161  ! P51195 porphyra purpurea.
SwissProt_40:PHAC_PORPU +    44.00  115.84  58.09      186  ! Q9n2j0 felis silvestris ca
SwissProt_40:INB_FELCA +    44.00  114.72  58.04      352  ! P47793 brachydanio rerio
SwissProt_40:WNT4_BRARE -    44.00  109.78  57.80

seq_name: SwissProt_40:INB_HUMAN

seq_documentation_block:
ID   INB_HUMAN       STANDARD;       PRT:   187 AA.
AC   P01574;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Interferon beta precursor (IFN-beta) (Fibroblast interferon).
GN   IFNB1 OR IFNB OR IFB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81198952; PubMed=6164984;
RA   Lawn R.M., Adelman J., Franke A.E., Houck C.M., Gross M.,
RA   Najarian R., Goeddel D.V.;
RT   "Human fibroblast interferon gene lacks introns.";
RL   Nucleic Acids Res. 9:1045-1052(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Ohno S., Taniguchi T.;
RT   "Structure of a chromosomal gene for human interferon beta.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:5305-5309(1981).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81005005; PubMed=6157601;
RA   Taniguchi T., Ohno S., Fujii-Kuriyama Y., Muramatsu M.;
RT   "The nucleotide sequence of human fibroblast interferon cDNA.";
RL   Gene 10:11-15(1980).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 80254542; PubMed=6157094;
RA   Derynck R., Content J., Declercq E., Volckaert G., Tavernier J.,
RA   Devos R., Fiers W.;
RT   "Isolation and structure of a human fibroblast interferon gene.";
RL   Nature 285:542-547(1980).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81053720; PubMed=6159580;
RA   Houghton M., Eaton M.A.W., Stewart A.G., Smith J.C., Doel S.M.,
RA   Catlin G.H., Lewis H.M., Patel T.P., Emtage J.S., Carey N.H.,
RA   Porter A.G.;
RT   "The complete amino acid sequence of human fibroblast interferon as
RT   deduced using synthetic oligodeoxyribonucleotide primers of reverse
RT   transcriptase.";
RL   Nucleic Acids Res. 8:2885-2894(1980).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81053752; PubMed=6159584;
RA   Goeddel D.V., Shepard H.M., Yelverton E., Leung D., Crea R., Sloma A.,
RA   Pestka S.;
RT   "Synthesis of human fibroblast interferon by E. coli.";
RL   Nucleic Acids Res. 8:4057-4074(1980).
RN   [7]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80063565; PubMed=2414376;
RA   May L.T., Sehgal P.B.;
RT   "On the relationship between human interferon alpha 1 and beta 1
RT   genes.";
RL   J. Interferon Res. 5:521-526(1985).
RN   [8]
RP   SEQUENCE OF 1-68 FROM N.A.
RX   MEDLINE=81053854; PubMed=6159597;
RA   Houghton M., Stewart A.G., Doel S.M., Emtage J.S., Eaton M.A.W.,
```

```
RA   Smith J.C., Patel T.P., Lewis H.M., Porter A.G., Birch J.R.,
RA   Cartwright T., Carey N.H.;
RT   "The amino-terminal sequence of human fibroblast interferon as
RT   deduced from reverse transcripts obtained using synthetic
RT   oligodeoxynucleotide primers.";
RL   Nucleic Acids Res. 8:4113-4131(1980).
RN   [5]
RP   DISULFID BOND.
RX   MEDLINE=81120983; PubMed=6162107;
RA   Wetzel R.;
RT   "Assignment of the disulphide bonds of leukocyte interferon.";
RL   Nature 289:606-607(1981).
RN   [6]
RP   SEQUENCE FROM N.A. (VARIANT CLONE PF526).
RX   MEDLINE=84080883; PubMed=6171745;
RA   Shepard H.M., Lewis G., Stebbing N., Goeddel D.V.;
RT   "A single amino acid change in IFN-beta abolishes its antiviral
RT   activity.";
RL   Nature 294:563-565(1981).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=98003481; PubMed=9342320;
RA   Karpusas M., Nolte M., Benton C.B., Meier W., Lipscomb W.N., Goelz S.;
RT   "The crystal structure of human interferon beta at 2.2-A resolution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:11813-11818(1997).
CC   -!- FUNCTION: HAS ANTIVIRAL, ANTIBACTERIAL AND ANTICANCER ACTIVITIES.
CC   -!- SUBUNIT: MONOMER.
CC   -!- PHARMACEUTICAL: Available under the names Avonex (Biogen),
CC       Betaseron (Berlex) and Rebif (Serono). Used in the treatment of
CC       multiple sclerosis (MS). Betaseron is a slightly modified form
CC       of IFNB1 with two residue substitutions.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   -!- DATABASE: NAME=Avonex; NOTE=Clinical information on Avonex;
CC       WWW "http://www.avonex.com/";
CC   -!- DATABASE: NAME=Betaseron; NOTE=Clinical information on Betaseron;
CC       WWW "http://www.betaseron.com/";
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; V00594; AAA52705.1; -.
DR   EMBL; V00545; AAA52706.1; -.
DR   EMBL; V00534; CAA23867.1; -.
DR   EMBL; V00547; CAA23868.1; -.
DR   EMBL; M28622; AAA36040.1; -.
DR   PIR; A00001; IFHU.
DR   HSSP; P01574; 1AU1.
DR   MIM; 147640; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000568; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Glycoprotein; Antiviral; Signal; Pharmaceutical;
KW   3D-structure.
FT   SIGNAL       1     21
FT   CHAIN       22    187       INTERFERON BETA.
FT   CARBOHYD     80     80       N-LINKED (GLCNAC...).
FT   DISULFID     36    148
FT   VARIANT     140    140       C     Y (IN CLONE PF526; LOSS OF ABILITY
FT                                 TO FORM THE ESSENTIAL DISULFIDE BOND;
FT                                 LOSS OF ANTIVIRAL ACTIVITY).
FT                                 /FTId=VAR_004516.
SQ   SEQUENCE   187 AA;  22204 MW;  0B01E4D408772A0C CRC64;
```

```
alignment_scores:
            Quality:   89.00              Length:     22
              Ratio:   4.944                Gaps:      0
 Percent Similarity:   81.818   Percent Identity: 77.273

alignment_block:
 US 09-832-658-7 x INB_HUMAN   ..

 Align seq 1/1   to: INB_HUMAN  from: 1   to: 187

       21 ATGCCTTACCCCGCTCTTCGGAGCCCTACAACCTTCTAGCAAATTTCAGTG 70
          |||:::|||      ||||||   ||||||   |||||||||||||||||
       22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCys 68

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       38 sGlnLysLeuLeuTrp 43

seq_name: SwissProt_40:DREB_HUMAN

seq_documentation_block:
ID   DREB_HUMAN      STANDARD;      PRT:   649 AA.
AC   Q16643;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-OCT-2001 (Rel. 40, Last annotation update)
DE   Drebrin E.
GN   DBN1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Osteoblast;
RA   Fisher L.W., McBride O.W., Filpula D., Ibaraki K., Young M.F.;
RT   "Human drebrin, cDNA sequence, mRNA tissue distribution and
RT   chromosomal localization.";
RL   Neurosci. Res. Commun. 14:35-42(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=94030046; PubMed=8216329;
RA   [illegible] M., [illegible] T., Mio [illegible] S., Shimizu N., Soya S., Oyama F.;
RT   "Molecular cloning of cDNA encoding human drebrin E and chromosomal
RT   mapping of its gene.";
RL   Biochem. Biophys. Res. Commun. 196:468-472(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: DREBRINS MIGHT PLAY SOME ROLE IN CELL MIGRATION,
CC       EXTENSION OF NEURONAL PROCESSES AND PLASTICITY OF DENDRITES,
CC       RESPECTIVELY. BINDS ACTIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: BRAIN NEURONS. ALSO FOUND IN THE HEART,
CC       PLACENTA, SKELETAL MUSCLE, KIDNEY AND PANCREAS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U00802; AAA16256.1; -.
DR   EMBL; D17530; BAA04480.1; -.
DR   EMBL; BC000283; AAH00283.1; -.
DR   MIM; 126660; -.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
```

```
DR   SMART; SM00102; ADF; 1.
KW   Actin binding; Brain; Neurone.
SQ   SEQUENCE   649 AA;  71425 MW;  A17201705F0242D88 CRC64;


alignment_scores:
            Quality:   49.50              Length:      29
              Ratio:   2.152                Gaps:       1
 Percent Similarity:   79.310   Percent Identity: 41.379

alignment_block:
 US-09-832-658-7 x DREB_HUMAN    ..

 Align seq 1/1  to: DREB_HUMAN  from: 1  to: 649

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   |||::::|||   ::::::   |||::::||||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 53

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTCTGG 86
            |||::::::|||:::   ||||||:::::::::::
       53 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 65

Seq_name: SwissProt_40:DREB_RAT

Seq_documentation_block:
ID   DREB_RAT        STANDARD;       PRT:   707 AA.
AC   Q07266;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Drebrin A.
GN   DBN1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN WISTAR; TISSUE=Brain, and Hippocampus;
RX   MEDLINE=92305233; PubMed=1611026;
RA   Shirao T., Obata N., Obata K.;
RT   "Cloning of drebrin A and induction of neurite-like processes in
RT   drebrin-transfected cells.";
RL   NeuroReport 3:109-112(1992).
CC   -!- FUNCTION: DREBRINS MIGHT PLAY SOME ROLE IN CELL MIGRATION,
CC       EXTENSION OF NEURONAL PROCESSES AND PLASTICITY OF DENDRITES,
CC       RESPECTIVELY. BINDS ACTIN.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: MULTIPLE FORMS OF DREBRINS RESULT FROM
CC       ALTERNATIVE SPLICING OF THE SINGLE DREBRIN GENE DURING NEURAL
CC       DEVELOPMENT.
CC   -!- TISSUE SPECIFICITY: BRAIN NEURONS.
CC   -!- MISCELLANEOUS: DREBRINS ARE CLASSIFIED INTO TWO FORMS OF THE
CC       EMBRYONIC TYPE (E1 AND E2) AND ONE FORM OF THE ADULT TYPE (A).  THE
CC       TIME COURSE OF THEIR APPEARANCE ARE DIFFERENT FROM EACH OTHER.
CC       THEIR STRUCTURES ARE CLOSELY RELATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   ---------------------------------------------------------------------
DR   EMBL; X59267; CAA41957.1; -.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
DR   SMART; SM00102; ADF; 1.
KW   Actin binding; Brain; Neurone; Alternative splicing.
SQ   SEQUENCE   707 AA;  77471 MW;  B5279BF6EF7880AA CRC64;


alignment_scores:
            Quality:   49.50              Length:      29
              Ratio:   2.152                Gaps:       1
 Percent Similarity:   79.310   Percent Identity: 41.379

alignment_block:
 US-09-832-658-7 x DREB_RAT    ..

 Align seq 1/1  to: DREB_RAT  from: 1  to: 707

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   |||::::|||   ::::::   |||::::||||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 53

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTCTGG 86
            |||::::::|||:::   ||||||:::::::::::
       53 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 65

seq_name: SwissProt_40:PTHP_CHLTR

seq_documentation_block:
ID   PTHP_CHLTR      STANDARD;       PRT:   109 AA.
AC   O84341;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Phosphocarrier protein HPr (Histidine-containing protein).
GN   PTSH OR CT337.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/CX;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE DEPENDENT
CC       SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC       TRANSPORT SYSTEM. THE PHOSPHORYL GROUP FROM PHOSPHOENOLPYRUVATE
CC       (PEP) IS TRANSFERRED TO THE PHOSPHORYL CARRIER PROTEIN HPR BY
CC       ENZYME I. PHOSPHO-HPR THEN TRANSFERS IT TO THE PERMEASE (ENZYMES
CC       II/III). HPR IS COMMON TO ALL PTS (BY SIMILARITY).
CC   -!- ENZYME REGULATION: PHOSPHORYLATION ON SER-67 INHIBITS THE
CC       PHOSPHORYL TRANSFER FROM ENZYME I TO HPR (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE HPR FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001306; AAC67932.1; -.
DR   HSSP; P08877; 2HPR.
DR   InterPro; IPR000032; HPr_protein.
DR   InterPro; IPR001020; PTS_HPr_his.
DR   InterPro; IPR002114; PTS_HPr_ser.
DR   Pfam; PF00381; PTS-HPr; 1.
DR   PROSITE; PS00369; PTS_HPR_HIS; FALSE_NEG.
DR   PROSITE; PS00589; PTS_HPR_SER; 1.
KW   Phosphotransferase system; Sugar transport; Phosphorylation;
KW   Complete proteome.
FT   MOD_RES      36      36       PHOSPHORYLATION (BY ENZYME I) (BY
FT                                 SIMILARITY).
```

```
CC   MOL_REG          1:1       2:2        PHOSPHORYLATION (BY SIMILARITY).
SQ   SEQUENCE   109 AA;  11997 MW;  0315273884760B6CDEF CRC64;


Alignment Scores:
            Quality:   44.00              Length:      19
              Ratio:   2.267               Gaps:        0
 Percent Similarity:   78.947   Percent Identity: 47.368

Alignment block:
US-09-832-658-7 x EYA4_CHICK  ..

 Align seq 1/1  to: EYA4_CHICK  from: 1  to: 109

       61 CCCCTCTTCGGAGCCCTACAAGCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGT 80
          :|:|:::::    ::::::|   |||:::|:|::::::|||:::||
       16 ProLeuMetGluAspGlyGluLeuSerAlaValTheIleArgAsnSer Cys

       81 CTTTCC 86
          :  :||
       42 LeuSer 47

Search completed: SwissProt_40:EYA4_CHICK

Sequence documentation block:
ID   EYA4_CHICK      STANDARD;      PRT:   109 AA.
AC   Q90YB0;
DT   30-MAY-2001 (Rel. 39, Created)
DT   30-MAY-2001 (Rel. 39, Last sequence update)
DT   30-MAY-2001 (Rel. 39, Last annotation update)
DE   Eyes absent homolog 4 (Fragment).
GN   EYA4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=20104912; PubMed=9887327;
RA   Borsani G., DeGrandi A., Ballabio A., Bulfone A., Bernard L.,
RA   Banfi S., Gattuso C., Mariani M., Dixon M., Donnai D., Metcalfe K.,
RA   Winter R., Robertson M., Axton R., Brown A., van Heyningen V.,
RA   Hanson I.;
RT   "EYA4, a novel vertebrate gene related to Drosophila eyes absent.";
RL   Hum. Mol. Genet. 8:65-74(1999).
CC   -!- FUNCTION: MAY BE INVOLVED IN DEVELOPMENT OF THE EYE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EYA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF096004; AAC98424.1; -.
KW   Developmental protein; Multigene family.
FT   NON_TER       1       1
FT   NON_TER     109     109
SQ   SEQUENCE   109 AA;  12439 MW;  9E6B04E6FB0D679D CRC64;


Alignment Scores:
            Quality:   48.00              Length:      27
              Ratio:   2.694               Gaps:        1
 Percent Similarity:   66.667   Percent Identity: 44.444

Alignment block:
US-09-832-658-7 x EYA4_CHICK  ..
```

```
 Align seq 1/1  to: EYA4_CHICK  from: 1  to: 119

        4 GACGCGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCT 53
          |||:::::          |||||||||       |||          |||
       41 GluAlaLeuThrAspSerTrpLeuThrAsnAlaLeu.....LysSe 54

       54 TCTAGCAATTTTCAGTGTCAGAAGCTCCTGT 84
          |||:::|||:::|||:::||||||::|||
       54 rLeuSerIleIleSerThrArgSerAsnCys 64

Seq_name: SwissProt_40:WNT2_MOUSE

Sequence_documentation_block:
ID   WNT2_MOUSE      STANDARD;      PRT:   360 AA.
AC   P21552;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   WNT-2 protein precursor (IRP protein) (INT-1 related protein).
GN   WNT2 OR WNT 2 OR IRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90126394; PubMed=2693011;
RA   McMahon J.A., McMahon A.P.;
RT   "Nucleotide sequence, chromosomal localization and developmental
RT   expression of the mouse int-1-related gene.";
RL   Development 107:643-650(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91122634; PubMed=2279700;
RA   Gavin B.J., McMahon J.A., McMahon A.P.;
RT   "Expression of multiple novel Wnt-1/int-1-related genes during fetal
RT   and adult mouse development.";
RL   Genes Dev. 4:2319-2332(1990).
CC   -!- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC       TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A
CC       SIGNALING MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE
CC       REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL
CC       DIAMETERS.
CC   -!- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC       extracellular matrix.
CC   -!- TISSUE SPECIFICITY: IN EMBRYOS IN THE DEVELOPING ALLANTOIS,
CC       PERICARDIUM HEART, AND VENTRAL-LATERAL MESODERM, IN ADULTS IN
CC       LUNG, BRAIN, HEART AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
DR   PIR; B36470; B36470.
DR   PIR; A43558; A43558.
DR   MGD; MGI:98954; Wnt2.
DR   InterPro; IPR000970; Wnt1.
DR   Pfam; PF00110; wnt; 1.
DR   PRINTS; PR01349; WNTPROTEIN.
DR   SMART; SM00097; WNT1; 1.
DR   PROSITE; PS00246; WNT1; 1.
KW   Developmental protein; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    360       WNT-2 PROTEIN.
FT   CARBOHYD    295    295       N-LINKED (GLCNAC...) (POTENTIAL)
SQ   SEQUENCE   360 AA;  40496 MW;  5EC265FD3815EF1D CRC64;


alignment_scores:
            Quality:   48.00              Length:      25
              Ratio:   2.824               Gaps:        0
 Percent Similarity:   68.000   Percent Identity: 40.000

alignment_block:
 US-09-832-658-7/rev x WNT2_MOUSE  ..
```

```
Align seq 1/1  to: WNT2_MOUSE  from: 1  to: 360

      87 ACCACAGGAGCTCTGACACTGAAAATTGCTAGAAGCTTGTAGGGCTCCA 38
         ||||||||||::|||:::          |||   ::::::   ||||||
      42 AlaThrGlyGlySerSerArgValMetCysAspAsnValProGlyLeuVa 48

      37 AGAGGGGCGTAAGCCATCTTGTCAT 13
         :::|||::::::       ||||||
      48 SerArgGlnArgGlnLeuCysHis 56

Seq_name: SwissProt_10:PLSB_MOUSE

seq_documentation_block:
ID   PLSB_MOUSE      STANDARD;      PRT;   827 AA.
AC   Q61586;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE   (EC 2.3.1.15) (GPAT) (P90).
GN   GPAM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92084678; PubMed=1721057;
RA   Shin D.-H., Paulauskis J.D., Moustaid N., Sul H.S.;
RT   "Transcriptional regulation of p90 with sequence homology to
RT   Escherichia coli glycerol-3-phosphate acyltransferase.";
RL   J. Biol. Chem. 266:23834-23839(1991).
CC   -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC       acyl-sn-glycerol 3-phosphate.
CC   -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC       ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       outer membrane (By similarity).
CC   -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN LIVER, INTERMEDIATE LEVELS
CC       IN MUSCLE AND KIDNEY, AND LOWEST LEVELS IN LUNG AND BRAIN
CC   -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M77003; AAA37647.1; -.
DR   MGD; MGI:109162; Gpam.
DR   InterPro; IPR002123; Acyltransferase.
DR   Pfam; PF01553; Acyltransferase; 1.
KW   Phospholipid biosynthesis; Transferase; Acyltransferase;
KW   Transmembrane; Mitochondrion; Transit peptide.
FT   TRANSIT        1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN          ?    827       GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE
FT   DOMAIN         ?    471       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM     472    494       POTENTIAL.
FT   DOMAIN       495    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     575    593       POTENTIAL.
FT   DOMAIN       594    827       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL)
SQ   SEQUENCE   827 AA;  93689 MW;  A64EDE697BD6b4B3 CRC64;


 alignment_scores:
              Quality:    48.00              Length:      20
                Ratio:     3.429                Gaps:       0
  Percent Similarity:    70.000   Percent Identity: 55.000

 alignment_block:
```

```
US-09-832-658-7 x PLSB_MOUSE  ..

Align seq 1/1  to: PLSB_MOUSE  from: 1  to: 827

       7 ACGATCATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCT 56
         ||:::::::  |||   |||   ||||||||||||   ||||::
     718 ThrPheLeuGlnArgLeuLeuGlyProLeuLeuGluAlaPheSerSerAl 744

      57 AGCAATTTTC 66
         ||||||||||
     734 aAlaIlePhe 737

seq_name: SwissProt_40:PLSB_HUMAN

seq_documentation_block:
ID   PLSB_HUMAN      STANDARD;      PRT;   828 AA.
AC   Q9HCL2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE   (EC 2.3.1.15) (GPAT).
GN   GPAM OR KIAA1560.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ashwell R.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 167-828 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20450683; PubMed=10997877;
RA   Nagase T., Kikuno R., Nakayama M., Hirosawa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.
RT   XVIII. The complete sequences of 100 new cDNA clones from brain which
RT   code for large proteins in vitro.";
RL   DNA Res. 7:273-281(2000).
CC   -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC       acyl-sn-glycerol 3-phosphate.
CC   -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC       ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       outer membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL391986; -; NOT_ANNOTATED_CDS.
DR   EMBL; AB046780; BAB13386.1; -.
DR   MIM; 602395; -.
DR   InterPro; IPR002123; Acyltransferase.
DR   Pfam; PF01553; Acyltransferase; 1.
KW   Phospholipid biosynthesis; Transferase; Acyltransferase;
KW   Transmembrane; Mitochondrion; Transit peptide.
FT   TRANSIT        1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN          ?    828       GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
FT   DOMAIN         ?    471       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM     472    494       POTENTIAL.
FT   DOMAIN       495    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     575    593       POTENTIAL.
FT   DOMAIN       594    828       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
SQ   SEQUENCE   828 AA;  93836 MW;  E5CD321A23D0B65B CRC64;
```

```
alignment_scores:
            Quality:   48.00              Length:      20
              Ratio:   3.429               Gaps:        0
 Percent Similarity:   70.000   Percent Identity: 55.000

alignment_block:
 US-09-832-658-7 x PLSB_HUMAN   ..

 Align seq 1/1  to: PLSB_HUMAN  from: 1  to: 828

       7 ACGATGATGACAAGATGGCTTACGCTGCTCTTGGAGCCCTACAAGCTTCT 56
         ||:::::   |||   |||   ||||||||||||   ||:::
     719 ThrPheLeuGlnArgLeuLeuGlyProLeuLeuGluAlaTyrSerSerAl 735

      57 AGCAATTTTC 66
         ||||||||||
     735 aAlaIlePhe 738

seq_name: SwissProt_40:DREB_CHICK

seq_documentation_block:
ID   DREB_CHICK     STANDARD;      PRT:   652 AA.
AC   P18302; Q91358; Q91359;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Drebrins A, E1 and E2.
GN   DBN1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (FORMS E1 AND E2).
RC   TISSUE=Brain;
RX   MEDLINE=89089137; PubMed=3208110;
RA   Kojima N., Kato Y., Shirao T., Obata K.;
RT   "Nucleotide sequences of two embryonic drebrins, developmentally
RT   regulated brain proteins, and developmental change in their mRNAs.";
RL   Brain Res. 464:207-215(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (FORM A).
RX   MEDLINE=93368892; PubMed=8361332;
RA   Kojima N., Shirao T., Obata K.;
RT   "Molecular cloning of a developmentally regulated brain protein,
RT   chicken drebrin A and its expression by alternative splicing of the
RT   drebrin gene.";
RL   Brain Res. Mol. Brain Res. 19:101-114(1994).
CC   -!- FUNCTION: DREBRINS MIGHT PLAY SOME ROLE IN CELL MIGRATION,
CC       EXTENSION OF NEURONAL PROCESSES AND PLASTICITY OF DENDRITES,
CC       RESPECTIVELY.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: MULTIPLE FORMS OF DREBRINS RESULT FROM
CC       ALTERNATIVE SPLICING OF THE SINGLE DREBRIN GENE DURING NEURAL

FT   CHAIN         ?    828       GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
FT   DOMAIN        ?    471       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM    472    494       POTENTIAL.
FT   DOMAIN      495    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    575    593       POTENTIAL.
FT   DOMAIN      594    828       MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT   CONFLICT     47     47       C -> F (IN REF. 2).
FT   CONFLICT     85     85       S -> P (IN REF. 2).
FT   CONFLICT    300    300       I -> V (IN REF. 2).
FT   CONFLICT    331    331       L -> V (IN REF. 2).
FT   CONFLICT    472    474       ILE -> NLL (IN REF. 2).
FT   CONFLICT    497    497       R -> W (IN REF. 2).
FT   CONFLICT    602    606       SAGGL -> LPEP (IN REF. 2).
FT   CONFLICT    644    644       Q -> H (IN REF. 2).
FT   CONFLICT    744    744       G -> A (IN REF. 2).
SQ   SEQUENCE   828 AA;  93714 MW;  ACA4A087E7DEB12C CRC64;


alignment_scores:
            Quality:   48.00              Length:      20
              Ratio:   3.429               Gaps:        0
 Percent Similarity:   70.000   Percent Identity: 55.000

alignment_block:
 US-09-832-658-7 x PLSB_RAT   ..

 Align seq 1/1  to: PLSB_RAT  from: 1  to: 828

       7 ACGATGATGACAAGATGGCTTACGCTGCTCTTGGAGCCCTACAAGCTTCT 56
         ||:::::   |||   |||   ||||||||||||   ||:::
     719 ThrPheLeuGlnArgLeuLeuGlyProLeuLeuGluAlaTyrSerSerAl 735

      57 AGCAATTTTC 66
         ||||||||||
     735 aAlaIlePhe 738

seq_name: SwissProt_40:PLSB_RAT

seq_documentation_block:
ID   PLSB_RAT       STANDARD;      PRT:   828 AA.
AC   P97564; P97565; P97566; Q63449;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE   (EC 2.3.1.15) (GPAT).
GN   GPAM
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND ACTIVITY.
RC   TISSUE=Liver;
RX   MEDLINE=98276217; PubMed=9646428;
RA   Bhat B.G., Wang P., Kim J. H., Black L.M., Lewin T.M.,
RA   Prestwich G.D., Jr., Coleman R.A.;
RT   "Rat sn-glycerol-3-phosphate acyltransferase: molecular cloning and
RT   characterization of the cDNA and expressed protein.";
RL   Biochem. Biophys. Acta 1439:415-423(1999).
RN   [2]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TOPOLOGY.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX   MEDLINE=20409698; PubMed=10924502;
RA   Brilly V.S., Chakraborty T.R., Nikolov A.V., Morimoto T., Haldar D.,
RT   "Identification of two transmembrane regions and a cytosolic domain of
RT   rat mitochondrial glycerolphosphate acyltransferase.";
RL   J. Biol. Chem. 275:31668-31673(2000).
CC   -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol-3-phosphate = CoA + 1-
CC       acyl-sn-glycerol-3-phosphate.
CC   -!- PATHWAY: FIRST STEP IN DE NOVO TRIGLYCERIDE BIOSYNTHESIS. IT MAY
CC       ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       outer membrane.
CC   -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF021448; AAB71005.1; -.
DR   EMBL; U96771; AAB94170.2; ALT_INIT.
DR   InterPro; IPR002123; Acyltransferase.
DR   Pfam; PF01553; Acyltransferase; 1.
KW   Phospholipid biosynthesis; Transferase; Acyltransferase;
KW   Transmembrane; Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
```

```
CC       DEVELOPMENT.
CC   -!- TISSUE SPECIFICITY: BRAIN NEURONS.
CC   -!- MISCELLANEOUS: DEBRIN ARE CLASSIFIED INTO TWO FORMS OF THE
CC       EMBRYONIC TYPE (E1 AND E2) AND TWO FORM OF THE ADULT TYPE (A). THE
CC       TIME COURSE OF THEIR APPEARANCE ARE DIFFERENT FROM EACH OTHER.
CC       THEIR STRUCTURES ARE CLOSELY RELATED.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M36961; AAA48750.1; -.
DR   EMBL; S65296; AAB28012.1; -.
DR   EMBL; S65279; AAB28012.1; JOINED.
DR   EMBL; S65280; AAB28012.1; JOINED.
DR   EMBL; S65281; AAB28012.1; JOINED.
DR   EMBL; S65288; AAB28012.1; JOINED.
DR   EMBL; S65289; AAB28012.1; JOINED.
DR   EMBL; S65290; AAB28012.1; JOINED.
DR   EMBL; S65291; AAB28012.1; JOINED.
DR   EMBL; S65292; AAB28012.1; JOINED.
DR   EMBL; S65294; AAB28012.1; JOINED.
DR   EMBL; S65267; AAB28011.2; -.
DR   EMBL; S65295; AAB28013.1; -.
DR   PIR; A43776; A43776.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
DR   SMART; SM00102; ADF; 1.
KW   Actin-binding; Brain; Neurone; Alternative splicing.
FT   DOMAIN        47     50       POLY-GLY.
FT   DOMAIN       332    339       POLY-SER.
FT   DOMAIN       542    553       POLY-GLU.
FT   DOMAIN       643    646       POLY-GLU.
FT   VARSPLIC     317    362       MISSING (IN ISOFORM E1).
FT   VARSPLIC     527    527       MISSING (IN ISOFORM E2).
FT   CONFLICT     552    552       E -> EE (IN REF. 1).
SQ   SEQUENCE   652 AA;  71845 MW;  8EFF72CA5498A3E65 CRC64;


alignment_scores:
            Quality:   47.50              Length:     29
              Ratio:  2.065                Gaps:      1
Percent Similarity:  79.310   Percent Identity: 37.931

alignment_block:
 US-09-832-658-7 x DREB_CHICK

 Align seq 1/1  to: DREB_CHICK  from: 1  to: 652

       8 GGAGACGATGATGACAAGATGGCTTACGCCGCCTCTTGGAGCCC...CTACA 49
         |||::|||||||   |||::|||   ::::::   ||||::   |||::
      47 GlySerAspAspLeuLysSteuAlaAlaSerGlyGlyLeuLeuG: 53

      50 AGCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
         :   |||::::::::::::   ||||||||||||||
      54 nLeuSerGlyHisPheGluTleGlnLysValMetTyr 65

 Seq_name: SwissProt_40:IL1B_TRIVU

 Seq_documentation_block:
 ID   IL1B_TRIVU     STANDARD;      PRT:   269 AA.
 AC   Q9XS77;
 DT   01-MAR-2002 (Rel. 41, Created)
 DT   01-MAR-2002 (Rel. 41, Last sequence update)
 DT   01-MAR-2002 (Rel. 41, Last annotation update)
 DE   Interleukin-1 beta precursor (IL-1 beta).
 GN   IL1B.
 OS   Trichosurus vulpecula (Brush-tailed possum).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 OX   NCBI_TaxID=9337;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   MEDLINE  2022044; PubMed=10209203;
 RA   Wedlock D.N., Goh L.P., Parlane N.A., Buddle B.M.;
 RT   "Molecular cloning and physiological effects of brushtail possum
 RT   interleukin-1beta.";
 RL   Vet. Immunol. Immunopathol. 67:359-372(1999).
 CC   -!- FUNCTION: PRODUCED BY ACTIVATED MACROPHAGES, IL-1 STIMULATES
 CC       THYMOCYTE PROLIFERATION BY INDUCING IL-2 RELEASE, B-CELL
 CC       MATURATION & PROLIFERATION, & FIBROBLAST GROWTH FACTOR ACTIVITY.
 CC       IL-1 PROTEINS ARE INVOLVED IN THE INFLAMMATORY RESPONSE, BEING
 CC       IDENTIFIED AS ENDOGENOUS PYROGENS, AND ARE REPORTED TO STIMULATE
 CC       THE RELEASE OF PROSTAGLANDIN AND COLLAGENASE FROM SYNOVIAL CELLS.
 CC   -!- SUBUNIT: MONOMER.
 CC   -!- DOMAIN: THE SIMILARITY AMONG THE IL-1 PRECURSORS SUGGESTS THAT THE
 CC       AMINO ENDS OF THESE PROTEINS SERVE SOME AS YET UNDEFINED FUNCTION.
 CC   -!- MISCELLANEOUS: THE LACK OF A SPECIFIC HYDROPHOBIC SEGMENT IN THE
 CC       PRECURSOR SEQUENCE SUGGESTS THAT IL-1 IS RELEASED BY DAMAGED CELLS
 CC       OR IS SECRETED BY A MECHANISM DIFFERING FROM THAT USED FOR OTHER
 CC       SECRETORY PROTEINS.
 CC   -!- SIMILARITY: BELONGS TO THE IL-1 FAMILY.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; AF071539; AAD21871.1; -.
 DR   HSSP; P01584; 1HIB.
 DR   InterPro; IPR002348; IL1_HBGF.
 DR   InterPro; IPR000975; Interleukin_1.
 DR   InterPro; IPR001502; Interleukin_1_prop.
 DR   Pfam; PF00340; IL1; 1.
 DR   Pfam; PF02394; IL1_propep; 1.
 DR   PRINTS; PR00262; IL1HBGF.
 DR   ProDom; PD002536; Interleukin_1; 1.
 DR   SMART; SM00125; IL1; 1.
 DR   PROSITE; PS00254; INTERLEUKIN_1; 1.
 KW   Cytokine; Macrophage; Mitogen; Inflammatory response; Pyrogen.
 FT   PROPEP        1    112       BY SIMILARITY.
 FT   CHAIN       113    269       INTERLEUKIN-1 BETA.
 SQ   SEQUENCE   269 AA;  31141 MW;  307A1FE3B627D6E7 CRC64;


 alignment_scores:
             Quality:   47.00              Length:     28
               Ratio:  2.765                Gaps:      1
 Percent Similarity:  60.714   Percent Identity: 42.857

 alignment_block:
  US-09-832-658-7 x IL1B_TRIVU

  Align seq 1/1  to: IL1B_TRIVU  from: 1  to: 269

        6 GACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTC 55
          ||||||||||||:.....  |||  |||  |||      |||  ::
       17 AspAspAspAspArgGlnPheTyrGluAlaAspGlyProAsnGlnLysLy 33

       56 TAGCAATTTTCAG.............TGTCAGAAG 77
          ::::  ||||||              |||::|||
       33 sGlyTy:PheGlnTyrProAspLeuCysSerLys 44

 Seq_name: SwissProt_40:Y871_METJA

 Seq_documentation_block:
 ID   Y871_METJA     STANDARD;      PRT:   317 AA.
 AC   Q58281;
 DT   15-JUL-1998 (Rel. 36, Created)
```

```
DE   Hypothetical 78.8 kDa protein in HSP12-HXT10 intergenic region.
GN   YFL013C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN S288C / AB972;
RX   MEDLINE=95420292; PubMed=7670463;
RA   Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA   Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA   Yamazaki M., Tashiro H., Eki T.;
RT   "Analysis of the nucleotide sequence of chromosome VI from
RT   Saccharomyces cerevisiae.";
RL   Nat. Genet. 10:261-268(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / AB972;
RA   Barrell B.G., Churcher C., Rajandream M.A.;
RL   Submitted (SEP 1994) to the EMBL/GenBank/DDBJ databases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D50617; BAA09225.1; -.
DR   PIR; S44750; BAA09204.1; -.
DR   EMBL; Z46255; CAA86347.1; -.
DR   SGD; S0001881; YFL013C.
KW   Hypothetical protein.
SQ   SEQUENCE   692 AA;  78752 MW;  9C0084A3643IB241 CRC64;


alignment_scores:
            Quality:   47.00              Length:      24
              Ratio:   2.350               Gaps:        0
 Percent Similarity:   83.333   Percent Identity:  33.333

alignment_block:
 US-09-832-658-7 x YFB3_YEAST    ..

 Align seq 1/1  to: YFB3_YEAST  from: 1  to: 692

      12 GATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAA 61
         :::|||||||::::::::|||   |||   :::|||      ::::::::
     194 AsnAspLysLeuThrPheAlaGluLeuTyrValLeuIleIleAlaThrSe 210

      62 TTTTCAGTGTCAGAAGCTCCTG 83
         ::::::::|||::::|||::|||
     210 rPheLysCysSerLysIleLeu 217

Seq_name: SwissProt_40:ATM_HUMAN

seq_documentation_block:
ID   ATM_HUMAN       STANDARD;      PRT;  3056 AA.
AC   Q13315; O15429; Q93007; Q16551; Q12758; Q9NP02; Q9BCX7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Serine-protein kinase ATM (EC 2.7.1.37) (Ataxia telangiectasia
DE   mutated) (A-T, mutated).
GN   ATM.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
```

RX MEDLINE; 96154672; PubMed-8589678;
RA Savitsky K., Sfez S., Tagle D.A., Ziv Y., Sartiel A., Collins F.S.,
RA Shiloh Y., Rotman G.;
RT "The complete sequence of the coding region of the ATM gene reveals
RT similarity to cell cycle regulators in different species.";
RL Hum. Mol. Genet. 4:2025-2032(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97344327; PubMed-9199932;
RA Platzer M., Rotman G., Bauer D., Uziel T., Savitsky K., Bar-Shira A.,
RA Gilad S., Shiloh Y., Rosenthal A.;
RT "Ataxia-telangiectasia locus: sequence analysis of 184 kb of human
RT genomic DNA containing the entire ATM gene.";
RL Genome Res. 7:592-605(1997).
RN [3]
RP SEQUENCE OF 1-24 FROM N.A.
RX MEDLINE=97263793; PubMed-9108147;
RA [redacted] Savitsky K., Platzer M., Uziel T., Gilad S., Sartiel A., Rosenthal A.,
RA [redacted] Elroy-Stein O., Shiloh Y., Rotman G.;
RT "Ataxia telangiectasia: structural diversity of untranslated sequences
RT suggests complex post-transcriptional regulation of ATM gene
RT expression.";
RL Nucleic Acids Res. 25:1679-1684(1997).
RN [4]
RP SEQUENCE OF 1-1369 FROM N.A., AND VARIANT AT 2546-SER--ILE-2548 DEL.
RX MEDLINE=98381449; PubMed-9789452;
RA Byrd P.J., McConville C.M., Cooper P., Parkhill J., Stankovic T.,
RA McGuire G.M., Thick J.A., Taylor A.M.R.;
RT "Mutations revealed by sequencing the 5' half of the gene for ataxia
RT telangiectasia.";
RL Hum. Mol. Genet. 5:145-149(1996).
RN [5]
RP SEQUENCE OF 1349-3056 FROM N.A., AND VARIANT ASN-3003.
RX MEDLINE=96105020; PubMed-8521492;
RA Rasio D., Negrini M., Croce C.M.;
RT "Genomic organization of the ATM locus involved in ataxia-
RT telangiectasia.";
RL Cancer Res. 55:6053-6057(1995).
RN [6]
RP SEQUENCE OF 1349-3056 FROM N.A., AND VARIANTS AT 2427-LEU-ARG-2428
RP DEL; 2546-SER-ILE-2548 DEL AND SER-2860 DEL.
RC TISSUE=Fibroblast;
RX MEDLINE=95312868; PubMed-7792600;
RA Savitsky K., Bar-Shira A., Gilad S., Rotman G., Ziv Y., Vanagaite L.,
RA Tagle D.A., Smith S., Uziel T., Sfez S., Ashkenazi M., Pecker I.,
RA Frydman M., Harnik R., Patanjali S.R., Simmons A., Clines G.A.,
RA [redacted] Sartiel A., Gatti R.A., Chessa L., Sanal O., Lavin M.F.,
RA [redacted] Jaspers N.G.J., Taylor A.M.R., Arlett C.F., Miki T., Weissman S.M.,
RA Lovett M., Collins F.S., Shiloh Y.;
RT "A single ataxia telangiectasia gene with a product similar to PI-3
RT kinase.";
RL Science 268:1749-1753(1995).
RN [7]
RP PARTIAL SEQUENCE FROM N.A., AND VARIANTS CYS-49, ARG-1054, PHE-1420,
RP ILE-2079 AND ALA-2287.
RX MEDLINE=96275738; PubMed-8665503;
RA Vorechovsky I., Rasio D., Luo L., Monaco C., Hammarstroem L.,
RA Webster A.D.B., Zaloudik J., Barbanti-Brodano G., James M.R.,
RA Russo G., Croce C.M., Negrini M.;
RT "The ATM gene and susceptibility to breast cancer: analysis of 38
RT breast tumors reveals no evidence for mutation.";
RL Cancer Res. 56:2726-2732(1996).
RN [8]
RP PHOSPHORYLATION.
RX MEDLINE=97126018; PubMed-8969240;
RA Chen G., Lee E.Y.-H.P.;
RT "The product of the ATM gene is a 370-kDa nuclear phosphoprotein.";
RL J. Biol. Chem. 271:33693-33697(1996).
RN [9]
RP SUBCELLULAR LOCATION.
RX MEDLINE=97203148; PubMed-9050866;
RA Brown K.D., Ziv Y., Sadanandan S.N., Chessa L., Collins F.S.,
RA Shiloh Y., Tagle D.A.;
RT "The ataxia-telangiectasia gene product, a constitutively expressed
RT nuclear protein that is not up-regulated following genome damage.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:1840-1845(1997).
RN [10]
RP SUBCELLULAR LOCATION, AND VARIANTS AT 2546-S-I-2548 DEL AND Y 2924.
RX MEDLINE=97294602; PubMed-9150358;
RA Watters D., Khanna K.K., Beamish H., Birrell G., Spring K., Kedar P.,
RA Gatei M., Stenzel D., Hobson K., Kozlov S., Zhang N., Farrell A.,
RA Ramsay J., Gatti R.A., Lavin M.F.;
RT "Cellular localisation of the ataxia-telangiectasia (ATM) gene product
RT and discrimination between mutated and normal forms.";
RL Oncogene 14:1911-1921(1997).
RN [11]
RP KINASE ACTIVITY.
RX MEDLINE=97141775; PubMed-8988033;
RA Jung M., Kondratyev A., Lee S.A., Dimtchev A., Dritschilo A.;
RT "ATM gene product phosphorylates I kappa B-alpha.";
RL Cancer Res. 57:24-27(1997).
RN [12]
RP C-ABL BINDING.
RX MEDLINE=97311400; PubMed-9168117;
RA Shafman T., Khanna K.K., Kedar P., Spring K., Kozlov S., Yen T.,
RA Hobson K., Gatei M., Zhang N., Watters D., Egerton M., Shiloh Y.,
RA Kharbanda S., Kufe D., Lavin M.F.;
RT "Interaction between ATM protein and c-Abl in response to DNA
RT damage.";
RL Nature 387:520-523(1997).
RN [13]
RP P53 BINDING, AND KINASE ACTIVITY.
RX MEDLINE=98052351; PubMed-9843217;
RA Khanna K.K., Keating K.E., Kozlov S., Scott S., Gatei M., Hobson K.,
RA Taya Y., Gabrielli B., Chan D., Lees-Miller S.P., Lavin M.F.;
RT "ATM associates with and phosphorylates p53: mapping the region of
RT interaction.";
RL Nat. Genet. 20:398-400(1998).
RN [14]
RP BETA-ADAPTIN BINDING.
RX MEDLINE=98374320; PubMed-9707615;
RA Lim D.-S., Kirsch D.G., Canman C.E., Ahn J.-H., Ziv Y., Newman L.S.,
RA Darnell R.B., Shiloh Y., Kastan M.B.;
RT "ATM binds to beta adaptin in cytoplasmic vesicles.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10146-10151(1998).
RN [15]
RP PHOSPHORYLATION OF P53.
RX MEDLINE=98404273; PubMed-9733514;
RA Banin S., Moyal L., Shieh S.-Y., Taya Y., Anderson C.W., Chessa L.,
RA Smorodinsky N.I., Prives C., Reiss Y., Shiloh Y., Ziv Y.;
RT "Enhanced phosphorylation of p53 by ATM in response to DNA damage.";
RL Science 281:1674-1677(1998).
RN [16]
RP PHOSPHORYLATION OF P53, AND MUTAGENESIS OF ASP-2870 AND ASN-2875.
RX MEDLINE=98404274; PubMed-9733515;
RA Canman C.E., Lim D.-S., Cimprich K.A., Taya Y., Tamai K.,
RA Sakaguchi K., Appella E., Kastan M.B., Siliciano J.D.;
RT "Activation of the ATM kinase by ionizing radiation and
RT phosphorylation of p53.";
RL Science 281:1677-1679(1998).
RN [17]
RP DNA BINDING.
RX MEDLINE=99432198; PubMed-10500142;
RA Smith G.C.M., Cary R.B., Lakin N.D., Hann B.C., Teo S.-H., Chen D.J.,
RA Jackson S.P.;
RT "Purification and DNA binding properties of the ataxia-telangiectasia
RT gene product ATM.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:11134-11139(1999).
RN [18]
RP PHOSPHORYLATION OF BRCA1.
RX MEDLINE=20018333; PubMed-10550055;
RA Cortez D., Wang Y., Qin J., Elledge S.J.;
RT "Requirement of ATM-dependent phosphorylation of brca1 in the DNA
RT damage response to double-strand breaks.";
RL Science 286:1162-1166(1999).
RN [19]

```
RT   IDENTIFICATION OF AIM AS MEMBER OF BASC.
RX   MEDLINE 22415492; PubMed 10784165;
RA   Wang Y., Cortez D., Yazdi P., Neff N., Elledge S.J., Qin J.;
RT   "BASC, a super complex of BRCA1-associated proteins involved in the
RT   recognition and repair of aberrant DNA structures.";
RL   Genes Dev. 14:927-939(2000).
RN   [20]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE 20220912; PubMed 10766245;
RA   Lim D., Kim S.T., Xu B., Maser R.S., Lin J., Petrini J.H.J.,
RA   Kastan M.B.;
RT   "ATM phosphorylates p95/nbs1 in an S-phase checkpoint pathway.";
RL   Nature 404:613-617(2000).
RN   [21]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE 22206495; PubMed 10865545;
RA   Wu X., Ranganathan V., Weisman D.S., Heine W.F., Ciccone D.N.,
RA   O'Neill T.B., Crick K.E., Pierce K.A., Lane W.S., Rathbun G.,
RA   Livingston D.M., Weaver D.T.;
RT   "ATM phosphorylation of Nijmegen breakage syndrome protein is required
RT   in a DNA damage response.";
RL   Nature 405:477-482(2000).
RN   [22]
RP   PHOSPHORYLATION OF NIBRIN.
RX   MEDLINE 22294391; PubMed 10962069;
RA   Gatei M., Young D., Cerosaletti K.M., Desai-Mehta A., Spring K.,
RA   Kozlov S., Lavin M.F., Gatti R.A., Concannon P., Khanna K.;
RT   "ATM-dependent phosphorylation of nibrin in response to radiation
RT   exposure.";
RL   Nat. Genet. 25:115-119(2000).
RN   [23]
RP   PHOSPHORYLATION OF C-JUN.
RX   MEDLINE 20365735; PubMed 10910365;
RA   Liu S., Hsu N.S.Y., Zhang L., Chen P.-L., Ziv Y., Shiloh Y.,
RA   Lee E.Y.-H.P., Lee W.-H.;
RT   "Functional link of BRCA1 and ataxia telangiectasia gene product in
RT   DNA damage response.";
RL   Nature 406:210-215(2000).
RN   [24]
RP   VARIANTS ALA-GLY-243; ASP-256 SER; ILE-256 DEL AND TYR-282.

Alignment scores:
            Quality:   47.00              Length:     21
              Ratio:  2.238                Gaps:      1
 Percent Similarity:  76.190   Percent Identity: 47.619

Alignment block
 US-09-832-658-7 x AIM_HUMAN   ..

 Align seq 1/1  to: AIM_HUMAN  from: 1  to: 3056

       17 CAAGATCCTTATGGTCTTCTGGAGAGCTACAAGCTTCTAGCAATTTTC 66
          : ||||||   ||||||   :::   |||::::::|||  |||   |||
    2794 GluAspTyrArgHisTyrSerAspTyrArgProAsnAspPheSerAlaPheG 2800

       67 ATTGTCAGAAG 77
          : |||||||||
    2801 IleTyrSerLys 2811

Seq name: SwissProt 4:EYA4_FUGRU

Seq documentation block:
ID   EYA4_FUGRU     STANDARD;      PRT;   119 AA.
AC   Q90HA1;
DT   30-MAY-2000 (Rel. 38, Created)
DT   30-MAY-2000 (Rel. 38, Last sequence update)
DT   15-MAY-2000 (Rel. 38, Last annotation update)
DE   Eyes absent homolog 4 (Fragment).
GN   EYA4.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 99105912; PubMed 9887327;
RA   Borsani G., DeGrandi A., Ballabio A., Bulfone A., Bernard L.,
RA   Banfi S., Gattuso C., Mariani M., Dixon M., Donnai D., Metcalfe K.,
RA   Winter R., Estivill X., Axton R., Brown A., van Heyningen V.,
RA   Hanson I.;
RT   "EYA4, a novel vertebrate gene related to Drosophila eyes absent.";
RL   Hum. Mol. Genet. 8:11-23(1999).
CC   -!- FUNCTION: MAY BE INVOLVED IN DEVELOPMENT OF THE EYE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE EYA FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ007998; CAA07821.1; -.
DR   EMBL; AJ007999; CAA07821.1; JOINED.
DR   EMBL; AJ008000; CAA07821.1; JOINED.
DR   EMBL; AJ008001; CAA07821.1; JOINED.
KW   Developmental protein; Multigene family.
FT   NON_TER       1      1
FT   NON_TER     119    119
SQ   SEQUENCE   119 AA;  13349 MW;  9BB3E01C33720904 CRC64;


alignment scores:
            Quality:   46.50              Length:     27
              Ratio:  2.735                Gaps:      1
 Percent Similarity:  62.963   Percent Identity: 44.444

alignment block:
 US-09-832-658-7 x EYA4_FUGRU   ..

 Align seq 1/1  to: EYA4_FUGRU  from: 1  to: 119

       4 GAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCT 53
         |||::::::          |||||||||       |||         |||
      41 GluAlaLeuThrAspSerTrpLeuThrHisAlaLeu.....LysSe 94

      54 TCTAGCAATTTTCAGTGTCAGAAGCTCCTGT 84
         |||::|||::|||   ||||||::|||
      54 rLeuSerIleIleSerSerArgSerAsnCys 64
```

```
OM of: US-09-832-658-7 to: SPTREMBL_19:*   out_format : pfs

Date: Oct 9, 2002  3:16 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1998-2000 Compugen Ltd.

Command line parameters:
  -MODEL frame=_n2p.model -DEV x11
  -Q /cgn2_1/ptodata/[illegible]/app_query.fasta_1.144
  -DB SPTREMBL_19 -QFMT=fastan -SUFFIX n2p.rspt -GAPOP=12.000
  -GAPEXT=4.000 -MINMATCH 0.100 -LOOPCL=0.000 -LOOPEXT=0.000
  -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
  -FGAPOP 6.000  FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
  -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
  -TRANS human40.cdi  LIST=45  DOCALIGN 200 -THR_SCORE EST
  -THR_MAX=100 -THR_MIN=0 -ALIGN 15 -MODE=LOCAL -OUTFMT=pis
  -[illegible]RM ext  HEAPSIZE 500 -MINLEN 0 -MAXLEN=2000000000
  -[illegible]ER US09832658_3CGN1_1_93  NCPU=6 -1CPU=3 -LONGLOG
  -[illegible]AV_TIMEOUT 120 -WARN_TIMEOUT 30 -NO_XLPXY -WAIT -THREADS=1

Search Information block:
Query: US-09-832-658-7
Query length: 87
Database: SPTREMBL_19:*
Database sequences: 562222
Database length: 172994929
Search time (sec): 34.360000


score_list:
Sequence                  Strd  Orig    ZScore    EScore Len  ! Documentation  ..
sp_mammal:O77812           +    84.00   230.19   9.5e-05   187  ! O77812 macaca fascicularis (cra
sp_human:Q15943            +    81.00   221.52    0.0003   187  ! Q15943 homo sapiens (human). in
sp_mammal:Q29412           +    53.00   140.64    9.34     196  ! Q29412 sus scrofa (pig). interf
sp_organelle:Q9XL23        +    53.00   133.17    9.86     459  ! Q9xl23 smithornis sharpei, nadh
sp_invertebrate:O17729     +    50.00   124.18   30.05     477  ! O17729 caenorhabditis elegans
sp_plant:O65341            +    50.00   122.75   30.36     567  ! O65341 saccharum robustum, solu
sp_plant:O65342            +    50.00   122.75   30.36     567  ! O65342 saccharum officinarum (s
sp_[illegible]:Q9QI21      +    50.00   119.96   30.98     795  ! Q9qpl1 gallid herpesvirus 1. co
sp_invertebrate:Q9VFU1     -    50.00   111.54   32.32    2284  ! Q9vfu1 drosophila melanogast
sp_invertebrate:Q9GK86     -    50.00   111.40   32.96    2240  ! Q9gr86 drosophila melanogast
sp_organelle:Q9B703        +    49.50   125.39   35.48     346  ! Q9b703 apteryx mantelli, nadh d
sp_human:Q9UIZ5            +    49.50   120.17   36.85     651  ! Q9uiz5 homo sapiens (human). hy
sp_rodent:Q9QXS5           +    49.50   120.05   36.88     660  ! Q9qxs5 mus musculus (mouse). dr
sp_rodent:Q922X1           +    49.50   120.05   36.88     660  ! Q922x1 mus musculus (mouse). dr
sp_rodent:O70205           +    49.50   120.04   36.88     661  ! O70205 rattus norvegicus (rat).
sp_rodent:Q9QXS6           +    49.50   119.50   37.02     706  ! Q9qxs6 mus musculus (mouse). dr
sp_bacteriap:Q97DB9        -    49.00   127.21   41.71     233  ! Q97db9 clostridium acetobutylic
sp_organelle:Q9LD56        +    49.00   123.90   42.72     348  ! Q9ld56 leptolebias minimus, nad
sp_plant:Q9SNR3            +    49.00   121.63   43.43     458  ! Q9snr3 oryza sativa (rice). est
sp_organelle:Q91GW3        +    48.50   122.50   51.40     346  ! Q91gw3 rhea americana (greater
sp_organelle:O79393        +    48.50   122.50   51.40     346  ! O79393 rhea americana (greater
sp_plant:Q93VH4            +    48.50   119.85   52.40     477  ! Q93vh4 oryza sativa (rice). put
sp_rodent:Q99MC2           -    48.00   133.26   56.65      79  ! Q99mc2 rattus norvegicus (rat)
sp_rodent:Q9JKQ8           -    48.00   125.75   59.81     196  ! Q9jkq8 mus musculus (mouse). wo
sp_human:Q9P2W0            +    48.00   124.18   60.49     237  ! Q9p2w0 homo sapiens (human). pe
sp_bacteria:Q9RGL8         +    48.00   123.74   60.68     250  ! Q9rgl8 abiotrophia defectiva. p
sp_archeap:O58254          +    48.00   123.67   60.71     252  ! O58254 pyrococcus horikoshii. h
sp_rodent:Q9CZW3           -    48.00   120.73   62.02     360  ! Q9czw3 mus musculus (mouse). zi
sp_rodent:Q9DBW4           +    48.00   119.24   62.69     431  ! Q9dbw4 mus musculus (mouse). 12
sp_bacteriap:Q98ER5        +    48.00   118.96   62.82     446  ! Q98er5 rhizobium loti (mesorhiz
sp_plant:Q9AJ38            +    48.00   118.65   62.96     463  ! Q9aj38 dunaliella salina, chlor
sp_virus:Q91B52            -    48.00   115.15   64.57     707  ! Q91b52 human cytomegalovirus. u
sp_plant:Q9FVK1            -    48.00   113.85   65.18     828  ! Q9fvk1 picea abies (norway spru
sp_invertebrate:Q9N3L7     -    48.00   112.57   65.78     966  ! Q9n3l7 caenorhabditis elegans
sp_organelle:Q9B6Z2        +    47.50   119.61   74.47     346  ! Q9b6z2 casuarius bennetti (dwar
sp_organelle:Q9B6V6        +    47.50   119.61   74.47     346  ! Q9b6v6 dromaius novae hollandia
sp_organelle:Q958B7        +    47.50   119.61   74.47     346  ! Q958b7 pterocnemia pennata (les
sp_organelle:Q957Z4        +    47.50   119.61   74.47     346  ! Q957z4 casuarius casuarius (aus
sp_plant:Q94D52            +    47.50   115.80   76.55     549  ! Q94d52 oryza sativa (rice). p07
sp_organelle:Q95GC1        +    47.50   113.99   77.56     683  ! Q95gc1 paspalum haumanii, nadh
sp_human:Q96TA0            +    47.50   113.40   77.89     734  ! Q96ta0 homo sapiens (human).
sp_bacteriap:Q98KL0        +    47.50   110.41   79.59    1053  ! Q98kl0 rhizobium loti (meso
sp_mammal:Q95L30           -    47.00   130.91   81.75      74  ! Q95l30 equus caballus (horse)
sp_mammal:Q95LB8           +    47.00   129.86   82.47      84  ! Q95lb8 sus scrofa (pig). inte
sp_organelle:P92521        +    47.00   126.85   84.19     121  ! P92521 arabidopsis thaliana

seq_name: sp_mammal:O77812

seq_documentation_block:
ID   O77812       PRELIMINARY;      PRT;   187 AA.
AC   O77812;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   INTERFERON BETA.
GN   IFN-BETA.
OS   Macaca fascicularis (crab-eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99156140; PubMed=10048895;
RA   Matheux F., Le Grand R., Rousseau V., De Maeyer E., Dormont D.;
RA   Lauret E.;
RT   "Macaque lymphocytes transduced by a constitutively expressed
RT   interferon beta gene display an enhanced resistance to SIVmac251
RT   infection.";
RL   Hum. Gene Ther. 10:429-429(1999).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; AJ011909; CAA09862.1; -.
DR   HSSP; P01574; 1AU1.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   187 AA;  22266 MW;  376AEF73332B7807 CRC64;


alignment_scores:
            Quality:   84.00              Length:     22
              Ratio:  4.667                Gaps:       0
 Percent Similarity:  81.818   Percent Identity:  72.727

alignment_block:
 US-09-832-658-7 x O77812   ..

 Align seq 1/1  to: O77812  from: 1  to: 187

       21 ATGGCTTACCCCGCTCTTGCAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||:::|||          ||||||   ||||||   ||||||:::|||||||
       22 MetSerIleAsnLeuLeuGlyPheLeuGlnArgSerSerSerPheGlnCy 48

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       38 sGlnLysLeuLeuTrp 43

seq_name: sp_human:Q15943

seq_documentation_block:
ID   Q15943       PRELIMINARY;      PRT;   187 AA.
AC   Q15943;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   INTERFERON-BETA PRECURSOR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 93005550; PubMed 8386692;
RA   Diaz M.O., Romine M.F., Devos R., Cheronnet M., Contreras R.E.,
RA   Chuvpilo S., Derijard W.M., Staunskens P., Inverniti J., Taya Y.,
RA   Content J.;
RT   "The human interferon' and human immune interferon genes and their
RT   expression in homologous and heterologous cells.";
RL   Pharm. Ernst. M. Soc. Exp. Biol., Bd Biol. Sci. 299:29-38(1992).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; M25460; AAA41702.1; -.
DR   HSSP; P05434; 1AU1.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; UNKNOWN_1.
KW   Antiviral; Cytokine; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    187       INTERFERON BETA.
SQ   SEQUENCE   187 AA;  22251 MW;  D8EE65B6948620C1 CRC64;


alignment_scores:
            Quality:   81.00              Length:      22
              Ratio:   4.765               Gaps:        3
 Percent Similarity:  77.273   Percent Identity: 72.727

alignment_block:
 US-09-832-658-7 x Q29412   ..

 Align seq 1/1  to: Q29412  from: 1  to: 187

       21 ATGGCTTACGCCGCTCTTGGAGCCTACAAGCTTCTAGCAATTTTCAGTG 70
          ||||:|||   :::|||          |||   ||||||||||::   ||
       22 MetSerTyrAspValLeuArgTyrGlnGlnArgSerSerAsnLeuAlaCy 68

       71 TCAGAAGCTCCTG 83
          |||||||||||||
       68 sGlnLysLeuLeu 42

seq_name: sp_organelle:Q9XL23

seq_documentation_block:
ID   Q9XL23      PRELIMINARY;      PRT;   459 AA.
AC   Q9XL23;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).
OS   Smithornis sharpei.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Eurylaimidae;
OC   Smithornis.
OX   NCBI_TaxID=81930;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 98393707; PubMed 9724766;
RA   Mindell D.P., Sorenson M.D., Dimcheff D.E.;
RT   "Multiple independent origins of mitochondrial gene order in birds.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:10693-10697(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Mindell D.P., Sorenson M.D., Dimcheff D.E., Hasegawa M., Ast J.C.,
RA   Yuri T.;
RT   "Interordinal relationships of birds and other reptiles based on whole
RT   mitochondrial genomes.";
RL   Syst. Biol. 48:138-152(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Sorenson M.D., Dimcheff D.E., Ast J.C., Yuri T., Mindell D.P.;
RT   "Complete mitochondrial DNA sequences for five birds and a turtle.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE   NAD(+) + UBIQUINOL.
DR   EMBL; AF090340; AAD32523.1; -.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q1_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
KW   Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ   SEQUENCE   459 AA;  51128 MW;  169A645C277C7ED1 CRC64;


alignment_scores:
```

```
            Quality:   53.00              Length:      17
              Ratio:   4.417               Gaps:       0
 Percent Similarity:   70.588   Percent Identity: 58.824

alignment_block:
 US-09-832-658-7 x Q9XL23    ..

 Align seq 1/1  to: Q9XL23  from: 1  to: 459

        7 ACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCT 56
          |||:::      :::|||| |||||||   |||   ||||||||||   ||
       46 ThrIleSerLeuGlnIrpLeuIlaProThrLeuTyrProTyrLysAsnLe 52

       57 A 57
          |
       52 u 52

 Name: sp_invertebrate:O17729

 seq_documentation_block:
 ID   O17729      PRELIMINARY;      PRT;   477 AA.
 AC   O17729;
 DT   01-JAN-1998 (TrEMBLrel. 05, Created)
 DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT   31-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   D1086.6 PROTEIN.
 GN   D1086.6.
 OS   Caenorhabditis elegans.
 OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC   Rhabditidae; Peloderinae; Caenorhabditis.
 OX   NCBI_TaxID=6239;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Smye R.;
 RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE 99069613; PubMed=9851916;
 RA   none;
 RT   "Genome sequence of the nematode C.elegans: A platform for
 RT   investigating biology.";
 RL   Science 282:2012-2018(1998).
 DR   EMBL; Z81491; CAB04023.2; -.
 SQ   SEQUENCE   477 AA;  54218 MW;  8268E004D61A231C CRC64;

 alignment_scores:
            Quality:   50.00              Length:      24
              Ratio:   2.941               Gaps:       0
 Percent Similarity:   70.833   Percent Identity: 45.833

 alignment_block:
 US-09-832-658-7 x O17729    .

 Align seq 1/1  to: O17729  from: 1  to: 477

        6 GACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTC 55
          |||:||      :::::::::   |||:||   |||          ||||::
      103 AspAspValTyrArgGlnCysAlaAlaAlaGlyGlyAsnTyrTyrAlaGl 119

       56 TAGCAATTTTCAGTGTCAGAAG 77
          ::::|:|||||||||:::|||
      119 yAlaAsnPheGlnCysSerLys 126

 seq_name: sp_plant:O65341

 seq_documentation_block:
 ID   O65341      PRELIMINARY;      PRT;   567 AA.
 AC   O65341;
 DT   01-AUG-1998 (TrEMBLrel. 07, Created)
 DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   SOLUBLE ACID INVERTASE (FRAGMENT).
 OS   Saccharum robustum.
 OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC   Panicoideae; Andropogoneae; Saccharum.
 OX   NCBI_TaxID=62334;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=CV. MOLOKAI 5829; TISSUE=STEM APEX;
 RA   Albert H.H., Zhu Y.J., Moore P.H.;
 RT   "Differential expression of soluble acid invertase (SAI) genes
 RT   correlates to differences in sucrose accumulation in sugarcane.";
 RL   Submitted (MAY 1998) to the EMBL/GenBank/DDBJ databases.
 CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA D
 CC       FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC   -!- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 DR   EMBL; AF062734; AAC16654.1; -.
 DR   InterPro; IPR001362; Glyco_hydro_32.
 DR   Pfam; PF00251; Glyco_hydro_32; 1.
 DR   PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
 KW   Glycoprotein; Glycosidase; Hydrolase.
 FT   NON_TER       1       1
 SQ   SEQUENCE   567 AA;  62651 MW;  92FF2231877ZTBB3 CRC64;

 alignment_scores:
            Quality:   50.00              Length:      20
              Ratio:   3.333               Gaps:       0
 Percent Similarity:   75.000   Percent Identity: 50.000

 alignment_block:
 US-09-832-658-7 x O65341    ..

 Align seq 1/1  to: O65341  from: 1  to: 567

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAAGC 52
          |||   ||||||||||:::      |||:::|||||||||      ::|||
      258 GlyMetAspAspAspArgHisAspTyrCysAlaLeuGlyArgTyrAspAl 274

       53 TTCTAGCAAT 62
          |:::::|||
      274 aAlaAlaAsn 277

 seq_name: sp_plant:O65342

 seq_documentation_block:
 ID   O65342      PRELIMINARY;      PRT;   567 AA.
 AC   O65342;
 DT   01-AUG-1998 (TrEMBLrel. 07, Created)
 DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   SOLUBLE ACID INVERTASE (FRAGMENT).
 OS   Saccharum officinarum (Sugarcane).
 OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC   Panicoideae; Andropogoneae; Saccharum.
 OX   NCBI_TaxID=4547;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=CV. LOUISIANA PURPLE; TISSUE=STEM APEX;
 RA   Albert H.H., Zhu Y.J., Moore P.H.;
 RT   "Differential expression of soluble acid invertase (SAI) genes
 RT   correlates to differences in sucrose accumulation in sugarcane.";
 RL   Submitted (MAY 1998) to the EMBL/GenBank/DDBJ databases.
 CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA D
 CC       FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.
 CC   -!- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 DR   EMBL; AF062735; AAC16655.1; -.
 DR   InterPro; IPR001362; Glyco_hydro_32.
 DR   Pfam; PF00251; Glyco_hydro_32; 1.
 DR   PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
 KW   Glycoprotein; Glycosidase; Hydrolase.
 FT   NON_TER       1       1
 SQ   SEQUENCE   567 AA;  62637 MW;  47A1FC84CFEB2AFB CRC64;
```

```
Alignment Scores:
            Quality:   [illegible]          Length:      [illegible]
              Ratio:   3.333                  Gaps:      [illegible]
Percent Similarity:   [illegible]  Percent Identity:  50.000

Matching Block:
SQ:  [illegible] x  [illegible]

 Query seq 1/1  [illegible]  from: 1  to: 562

       5 [illegible]  [illegible]
         [illegible]
     256 LysMetAsnAsnAsnArgHisAspTyrTyrAlaLeuGlyAlaTyrAspAl 274

      [illegible]
         [illegible]
     274 GlnAlaAlaGln 277

Seq Name: sp [illegible]

Documentation Block:
ID   GOGFII       PRELIMINARY;      PRT;   795 AA.
AC   GOGFII;
DT   01 MAY 2000 (IrEMBLrel. 14, Created)
DT   01 MAY 2000 (IrEMBLrel. 14, Last sequence update)
DT   01 MAY 2000 (IrEMBLrel. 14, Last annotation update)
DE   COMPONENT OF HELICASE PRIMASE COMPLEX.
GN   OLB.
OS   Gallid herpesvirus 2.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Varicella virus.
OX   NCBI_TaxID=10390;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96046[illegible]; PubMed=8466817;
RA   Fuchs W., Mettenleiter T.C.;
RT   "DNA sequence of the UL6 to UL20 genes of infectious laryngotracheitis
RT   virus and characterization of the UL10 gene product as a
RT   nonglycosylated membrane-associated virion protein.";
RL   J. Gen. Virol. 77:2159-2162(1996).
DR   EMBL; A[illegible]; AA[illegible].
KW   Helicase.
SQ   SEQUENCE   795 AA;  [illegible] MW;  [illegible] CRC64;


Alignment Scores:
            Quality:   50.000               Length:      17
              Ratio:   3.333                  Gaps:       0
Percent Similarity:   88.235  Percent Identity:  64.706

Matching Block:
SQ:  [illegible] x GOGFII   ..

 Alignment 1/1  to: GOGFII  from: 1  to: 795

      18 AACATCGCTTACCCCCCTCTTCCGACCCCTACAACCTTCTACCAATTTTCA 67
         ::::|||||:  |||:::  ::::|||||||::|||||||||||||||||
     1>1 ArgMetAlaThrAlaSerTyrAlaAlaLeuAsnAlaSerSerAsnPheGl 177

      68 G 68
         |
     178 L 177

Seq Name: sp [illegible]

Documentation Block:
ID   GOVEII       PRELIMINARY;      PRT;  2204 AA.
AC   GOVEII;
DT   [illegible] MAY 2000 (IrEMBLrel. 14, Created)
DT   [illegible] MAY 2000 (IrEMBLrel. 14, Last sequence update)
DT   [illegible] JUN 2001 (IrEMBLrel. 17, Last annotation update)
DE   PUTATIVE CADHERIN PRECURSOR (CG14900 PROTEIN).
GN   CG14900.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BERKELEY;
RX   MEDLINE 20196006; PubMed 10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
CC   -!- FUNCTION: MAY BE INVOLVED IN CELL ADHESION.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
DR   EMBL; AE003718; AAF55827.1; -.
DR   HSSP; P15116; 1NCJ.
DR   FlyBase; FBgn0038439; Cad89D.
DR   InterPro; IPR002126; Cadherin.
DR   Pfam; PF00028; cadherin; 9.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 12.
DR   PROSITE; PS00232; CADHERIN_1; 1.
DR   PROSITE; PS50268; CADHERIN_2; 12.
KW   Hypothetical protein; Cell adhesion; Glycoprotein; Calcium binding;
KW   Transmembrane; Signal; Repeat.
FT   SIGNAL        1      ?       POTENTIAL.
FT   CHAIN         ?   2204       PUTATIVE CADHERIN.
FT   DOMAIN        ?   1847       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1848   1868       POTENTIAL.
FT   DOMAIN     1869   2204       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42    143       CADHERIN 1.
FT   DOMAIN      152    259       CADHERIN 2.
FT   DOMAIN      289    475       CADHERIN 3.
FT   DOMAIN      384    492       CADHERIN 4.
FT   DOMAIN      502    589       CADHERIN 5.
```

```
FT   DOMAIN        783    841       CADHERIN 6.
FT   DOMAIN        900   1051       CADHERIN 7.
FT   DOMAIN       1143   1248       CADHERIN 8.
FT   DOMAIN       1257   1354       CADHERIN 9.
FT   DOMAIN       1383   1483       CADHERIN 10.
FT   DOMAIN       1506   1624       CADHERIN 11.
FT   DOMAIN       1635   1723       CADHERIN 12.
FT   CARBOHYD       78     78       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD       83     83       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      298    298       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      381    381       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      549    549       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      684    684       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      716    716       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      786    786       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      797    797       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      947    947       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      953    953       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      970    970       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     1219   1219       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     1282   1282       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     1450   1450       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     1493   1493       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     1520   1520       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   2204 AA;  246927 MW;  4C8DF541AD0CB150 CRC64;


alignment_scores:
                Quality:    50.00              Length:      27
                  Ratio:     2.500               Gaps:       0
  Percent Similarity:    74.074    Percent Identity: 40.741

alignment_block:
 US-09-832-658-7/rev x Q9VEU1    ..

 Align seq 1/1  to: Q9VEU1  from: 1  to: 2204

       83 CAGGAGCTTCTGACACTGAAAATTGCTAGAAGCTTGTAGGGCTCCAAGAG 34
          |||:::|||   :::|||||||||||      :::      ||||::   ||
     1052 GlnGlnLeuTyrAsnLeuLysIleIleAlaAlaAspThrGlyGlyLeuSe 1032

       33 CGGCGTAAGCCATCTTGTCATCATCGTCTCC 3
          ||||   :::|||||:::::::::|||:::
     1032 rGlyTyrAlaHisLeuThrValLeuValAla 1042

 Seq_name: sp_invertebrate:Q9GR86

 Seq_documentation_block:
 ID   Q9GR86       PRELIMINARY;      PRT;  2240 AA.
 AC   Q9GR86;
 DT   01-MAR-2001 (TrEMBLrel. 16, Created)
 DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   CAD89D.
 GN   CAD89D OR CG14900.
 OS   Drosophila melanogaster (Fruit fly).
 OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC   Ephydroidea; Drosophilidae; Drosophila.
 OX   NCBI_TaxID=7227;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Hirano S., Kimura H., Takeichi M., Uemura T.;
 RA   "A novel member of the Drosophila cadherin superfamily, Cad89D ";
 RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
 CC   -!- SIMILARITY: CONTAINS 12 CADHERIN DOMAINS.
 DR   EMBL; AB053260; BAB20634.1; -.
 DR   HSSP; P15116; 1NCJ.
 DR   FlyBase; FBgn0038439; Cad89D.
 DR   InterPro; IPR002126; Cadherin.
 DR   InterPro; IPR000504; RRM.
 DR   Pfam; PF00028; cadherin; 9.
 DR   PRINTS; PR00205; CADHERIN.
 DR   SMART; SM00112; CA; 12.
 DR   PROSITE; PS00232; CADHERIN_1; 1.
 DR   PROSITE; PS50268; CADHERIN_2; 12.
 DR   PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 KW   Calcium-binding; Cell adhesion; Glycoprotein.
 SQ   SEQUENCE   2240 AA;  250605 MW;  9487AF4AE7C81873 CRC64;


 alignment_scores:
                 Quality:    50.00              Length:      27
                   Ratio:     2.500               Gaps:       0
   Percent Similarity:    74.074    Percent Identity: 40.741

 alignment_block:
  US-09-832-658-7/rev x Q9GR86    ..

  Align seq 1/1  to: Q9GR86  from: 1  to: 2240

        83 CAGGAGCTTCTGACACTGAAAATTGCTAGAAGCTTGTAGGGCTCCAAGAG 34
           |||:::|||   :::|||||||||||      :::      ||||::   ||
      1052 GlnGlnLeuTyrAsnLeuLysIleIleAlaAlaAspThrGlyGlyLeuSe 1068

        33 CGGCGTAAGCCATCTTGTCATCATCGTCTCC 3
           ||||   :::|||||:::::::::|||:::
      1068 rGlyTyrAlaHisLeuThrValLeuValAla 1078

 seq_name: sp_organelle:Q9B703

 seq_documentation_block:
 ID   Q9B703       PRELIMINARY;      PRT;   346 AA.
 AC   Q9B703;
 DT   01-JUN-2001 (TrEMBLrel. 17, Created)
 DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   NADH DEHYDROGENASE SUBUNIT 2.
 OS   Apteryx mantelli.
 OG   Mitochondrion.
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Archosauria; Aves; Palaeognathae; Apterygiformes; Apterygidae;
 OC   Apteryx.
 OX   NCBI_TaxID=147461;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=21085666; PubMed=11217857;
 RA   Cooper A., Lalueza-Fox C., Anderson S., Rambaut A., Austin J.,
 RA   Ward R.;
 RT   "Complete mitochondrial genome sequences of two extinct moas clarify
 RT   ratite evolution.";
 RL   Nature 409:704-707(2001).
 CC   -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE   NAD(+) + UBIQUINOL.
 DR   EMBL; AY016010; AAK08536.1; -.
 DR   InterPro; IPR003917; NADHdh_oxdrdctse2.
 DR   InterPro; IPR001750; Oxidored_q1.
 DR   Pfam; PF00361; oxidored_q1; 1.
 DR   PRINTS; PR01436; NADHDHGNASE2.
 KW   Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ   SEQUENCE   346 AA;  38307 MW;  716CD31C19F7984C CRC64;


  alignment_scores:
                  Quality:    49.50              Length:      23
                    Ratio:     2.750               Gaps:       1
    Percent Similarity:    78.261    Percent Identity: 43.478

  alignment_block:
   US-09-832-658-7 x Q9B703    ..

   Align seq 1/1  to: Q9B703  from: 1  to: 346

         7 ACGATCATGACAAGATGGCTT...ACGCCGCTCTTGGAGCCCTACAAGCT 53
           |||||||||||   |||      ||||||:::|||:::   :::   ::
```

```
      279  IleMetMetThrSerIleThrLysThrProMetLeuAsnAlaIleLeuMe 245

       54  ICTAGCAATTTTCAGTGTC 72
           ||| ||||||||||||||||
      245  LeuThrLeuLeuSerLeu 251

Seq name: sp_human:Q9UEZ5

Seq documentation block:
ID   Q9UEZ5       PRELIMINARY;       PRT;   651 AA.
AC   Q9UEZ5;
DT   01-MAY-2000 (TrEMBLrel. 14, Created)
DT   01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 17, Last annotation update)
DE   HYPOTHETICAL 71.6 KDA PROTEIN.
GN   OKIAP1431004.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Wambutt R., Reinhardt D., Moreno H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (AUG 1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL110225; CAB53088.1; -.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
DR   SMART; SM00102; ADF; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   651 AA;  71764 MW;  8F27600AAAED5EC1 CRC64;


Alignment scores:
            Quality:   49.50              Length:      29
              Ratio:  2.152                Gaps:       1
 Percent Similarity:  79.310   Percent Identity: 41.379

Alignment block:
US-09-832-658-7 x Q9UEZ5   ..

 Align seq 1/1  to: Q9UEZ5  from: 1  to: 651

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   ||||::|||   ::::::   |||::::|||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 86

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCCTGTGG 86
             |||::::::::|||:::   |||||:::::::::
       87 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 123

Seq name: sp_rodent:Q922X1

Seq documentation block:
ID   Q922X1       PRELIMINARY;       PRT;   660 AA.
AC   Q922X1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DREBRIN 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC006714; AAH06714.1; -.
SQ   SEQUENCE   660 AA;  72416 MW;  9537DD52994C689 CRC64;


Alignment scores:
            Quality:   49.50              Length:      29
              Ratio:  2.152                Gaps:       1
 Percent Similarity:  79.310   Percent Identity: 41.379

Alignment block:
US-09-832-658-7 x Q922X1   ..

 Align seq 1/1  to: Q922X1  from: 1  to: 660

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   ||||::|||   ::::::   |||::::|||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 86

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCCTGTGG 86
             |||::::::::|||:::   |||||:::::::::
       87 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 123

Seq name: sp_rodent:O70205

Seq documentation block:
ID   O70205       PRELIMINARY;       PRT;   661 AA.
AC   O70205;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DR   EMBL; AF187148; AAF25190.1; -.
DR   MGD; MGI:1931838; Dbnl.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
DR   SMART; SM00102; ADF; 1.
SQ   SEQUENCE   660 AA;  72389 MW;  FB6F4B9F2A93BB27 CRC64;


Alignment scores:
            Quality:   49.50              Length:      29
              Ratio:  2.152                Gaps:       1
 Percent Similarity:  79.310   Percent Identity: 41.379

Alignment block:
US-09-832-658-7 x Q9QXS5   ..

 Align seq 1/1  to: Q9QXS5  from: 1  to: 660

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   ||||::|||   ::::::   |||::::|||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 86

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCCTGTGG 86
             |||::::::::|||:::   |||||:::::::::
       87 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 123

Seq name: sp_rodent:Q922X1

Seq documentation block:
ID   Q922X1       PRELIMINARY;       PRT;   660 AA.
AC   Q922X1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DREBRIN 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC006714; AAH06714.1; -.
SQ   SEQUENCE   660 AA;  72416 MW;  9537DD52994C689 CRC64;


Alignment scores:
            Quality:   49.50              Length:      29
              Ratio:  2.152                Gaps:       1
 Percent Similarity:  79.310   Percent Identity: 41.379

Alignment block:
US-09-832-658-7 x Q922X1   ..

 Align seq 1/1  to: Q922X1  from: 1  to: 660

        3 GGAGACGATGATGACAAGATGGCTTACGCCGCTCTTGGAGCCCTACAA.. 50
          |||::::||||||   ||||::|||   ::::::   |||::::|||||
       37 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 86

       51 .GCTTCTAGCAATTTTCAGTGTCAGAAGCTCCCTGTGG 86
             |||::::::::|||:::   |||||:::::::::
       87 uLeuSerGlyHisPheGluAsnGlnLysValMetTyr 123

Seq name: sp_rodent:O70205

Seq documentation block:
ID   O70205       PRELIMINARY;       PRT;   661 AA.
AC   O70205;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   DREBRIN F (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Eon Y., Kawai-hirai K., Xue Y., Shirao T.;
RT   "Study of Processes Formation in Drebrin cDNA Transfected fibroblast
RT   cells.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB015042; BAA28746.1; -.
DR   InterPro; IPR002108; Cofilin_ADF.
DR   Pfam; PF00241; cofilin_ADF; 1.
DR   SMART; SM00102; ADF; 1.
FT   NON_TER      661     661
SQ   SEQUENCE   661 AA;  72682 MW;  BF606FC889983495 CRC64;
```

```
Alignment_scores:
            Quality:    49.50              Length:      29
              Ratio:   2.152                Gaps:        1
 Percent Similarity:  79.310   Percent Identity: 41.379

Alignment_block:
 US-09-832-658-7 x o70205    ..

 Align seq 1/1:  to: o70205  from: 1  to: 661

        3 AGAGAAGATGATGACAACATGGCTTACGCCGCCTCTTGGAGCCCTACAA.. 50
          |||::::|||||  ||||::|||  ::::::  |||::|||||||
       47 GlySerAspAspLeuLysLeuAlaAlaSerGlyGluGlyGlyLeuGlnGl 53

       51 AGCTTCTAGCAATTTTCAGTGTCAGAAGCTCCTGTGG 86
             |||::::::|||:::  |:||||:::::::::
       53 uLeuSerGlyHisPheGlnAsnGlnLysValMetPyr 65
```

```
OM of: US-09-832-658-7 for: Issued Patents_AA:*  out_format : pir

Date: Oct 9, 2002  3:14 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd

Command line parameters:
 -MODEL trame7_n2p.model -DEV x11h
 -Q=/p2_5/ptodata/2/iaa/us-09-832-658-7.pir -SPOOL -Q_SEQ=/p2_5/ptodata/2/iaa/us-09-832-658-7_111358/tmp_1_1.rai_fmt.a ! 144
 -DB Issued_Patents_AA -QFMT -fastan -SUFFIX=n2p.rai
 -GAPOP 12.000 -GAPEXT 4.000 -MINMATCH 0.100 -LOOPCL=0.000
 -LOOPEXT 0.000 -QGAPOP=4.500 -QGAPEXT 0.050 -XGAPOP=10.000
 XGAPEXT 0.500  YGAPOP=6.000 -YGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT 0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=pct  THR_MAX=100  THR_MIN 0 -ALIGN 15 -MODE=LOCAL
 -OUTFMT=pir  NORM=ext  HEAPSIZE 500  MINLEN=0 -MAXLEN=2000000000
 -[illegible]ER US09832658_2CGN1_1_13 -NCPU 6 -ICPU=3 -LONGLOG
 -[illegible]EV_TIMEOUT 120  WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-832-658-7
Query length: 87
Database: Issued_Patents_AA:*
Database sequences: 231628
Database length: 24425594
Search time (sec): 16.460000

score_list:
Sequence                    Strd Orig    ZScore   EScore Len  ! Documentation ..
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-213-448-1 +    89.00  251.74  9.3e-07   166  !  S
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-362-453-16 +    89.00  251.74  9.3e-07   166  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-477-310A-1 +    89.00  251.74  9.3e-07   166  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-09-912-768-1 +    89.00  251.74  9.3e-07   166  !  S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-447-943A-2 +    89.00  251.74  9.3e-07   166  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:PCT-US95-04206-1 +    89.00  251.74  9.3e-07   166  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:5514567-4 +    89.00  251.74  9.3e-07   166  ! Paten
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-026,758-22 +    89.00  250.74  9.5e-07   187  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-09-912-768-3 +    89.00  250.74  9.5e-07   187  !  S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-296-903A-3 +    89.00  250.74  9.5e-07   187  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-426-930A-30 +    89.00  250.74  9.5e-07   187  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-002,122-9 +    89.00  250.74  9.5e-07   187  !  S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-296,935-7 +    89.00  250.74  9.5e-07   187  !  S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-296,936-7 +    89.00  250.74  9.5e-07   187  !  S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:5526459-2 +    89.00  250.74  9.5e-07   187  ! Paten
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:5514572-6 +    89.00  250.74  9.5e-07   187  ! Paten
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:5514567-3 +    89.00  250.74  9.5e-07   187  ! Paten
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-493-092-2 +    47.00  105.74  12.38     1708  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-508-836A-2 +    47.00  105.74  12.38     1708  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-508-836A-8 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-426-901A-3 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-074,266-2 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-442,224D-3 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-052,127-3 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-052-0140-3 +    47.00  100.59  13.39     3056  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-755,587-189 +    45.50  101.90  21.77     1589  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-367,782A-4 +    45.00  103.49  18.84      351  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-367,782A-2 +    45.00  103.76  21.52      351  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-474,027C-2 +    45.00  [illegible]  [illegible]  [illegible]  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-935,550-2 +    45.00   96.74  28.88     2516  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-414,416-3 +    45.00   94.75  28.74     3010  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-014,416-5 +    45.00   94.75  28.74     4011  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-841,271-12 +    44.50  107.57  27.99      597  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-751,767A-38 +    44.00  131.43  23.04       34  !
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-452-507-6    44.00  129.15  23.86       44  !  Se
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-188-799-87    44.00  126.12  24.98       62  !  S
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-440-103-16 +    44.00  114.13  30.45      269  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-440-542-16 +    44.00  114.13  30.45      269  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-[illegible]-16 +    44.00  114.13  30.45      269  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-09-440,210-16 +    44.00  114.13  30.45      269  !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-046-604-16 +    44.00  114.13  30.45      269  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-07-801-164A-4 +    43.00  112.81  42.88      199  !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-368,747-21 +    43.00  112.81  42.88      199  !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-087-943-3 +    43.00  107.28  46.65      372  !
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-460-806-1 +    43.00  106.59  47.14      402  !

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-213-448-1

seq_documentation_block:
; Sequence 1, Application US/08213448
; Patent No. 5545723
;  GENERAL INFORMATION:
;     APPLICANT:  Goelz, Susan E
;     APPLICANT:  Cate, Richard L
;     APPLICANT:  Pepinsky, Blake R
;     APPLICANT:  Chow, Pingchang E
;     TITLE OF INVENTION:  No. 5545723el Muteins of IFN-Beta
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  James F. Haley, Jr.
;       STREET:  Fish & Neave, 1251 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10020-1104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/213,448
;       FILING DATE:
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Haley Jr., James F
;       REGISTRATION NUMBER:  27,794
;       REFERENCE/DOCKET NUMBER:  8679
;     TELECOMMUNICATION INFORMATION
;       TELEPHONE:  (212) 596-9000
;       TELEFAX:  (212) 596-9090
;  INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  166 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-213-448-1


alignment scores:
            Quality:  89.00                Length:      22
              Ratio:  4.044                  Gaps:       0
  Percent Similarity:  81.818   Percent Identity:  77.273

alignment_block.
 US-09-832-658-7 x US-08-213-448-1  ..

 Align Seq 1/1  to  US-08-213-448  Len: 1 to 166

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGTAATTTTCAGTG 70
          |||...|||      ||||||  ||||||  ||||||||||||||||||
        1 MetSerTyrAsnLeuLeuGlyPheLeuGlnSerSerAsnPheGlnVal 47

       71 [illegible]  86
          ||||||||||||||||
       47 sGlnLysLeuLeuTrp 22

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-362-453-16
```

```
Seq_documentation_block:
; Sequence 16, Application US/08362453
; Patent No. 5984129
;  GENERAL INFORMATION:
;     APPLICANT:  FISH, Eleanor N.
;     TITLE OF INVENTION:  INTERFERON RECEPTOR BINDING PEPTIDES
;     NUMBER OF SEQUENCES:  17
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram
;       STREET:  655 Fifteenth Street N.W. Suite 330
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-5701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/362,453
;       FILING DATE:  06-JAN-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/609,739
;       APPLICATION NUMBER:  FILING DATE:  07-JUL-1992
;       APPLICATION NUMBER:  US 07/980,525
;       APPLICATION NUMBER:  FILING DATE:  20-NOV-1992
;       APPLICATION NUMBER:  US/08/083,279
;       APPLICATION NUMBER:  FILING DATE: 06-JUL-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kitts, Monica Chin
;       REGISTRATION NUMBER:  36,105
;       REFERENCE/DOCKET NUMBER:  PC08-4017
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 638-5601
;       TELEFAX:  (202) 638-4810
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  166 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
US 09 832 658 7

Alignment Scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:        0
 Percent Similarity:  81.818   Percent Identity: 77.273

Alignment Block:
 US-09-832-658-7 x US-08-362-453-16    ..

 Align seq 1/1  to: US 08 362 453 16  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||::|||      ||||||  ||||||  ||||||||||||||||||
        1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

Seq_name: /usr/u5/pl/Data/Output/A_7Mb.pep 'S' 1-422 vi.A 1

Seq_documentation_block:
; Sequence 1, Application US/08477410A
; Patent No. 5814480
;  GENERAL INFORMATION:
;     APPLICANT:  Dorin, Glenn
;     APPLICANT:  McAvoy, Patrick J.
;     APPLICANT:  Wood, Kathleen M.
;     TITLE OF INVENTION:  Bacterial Production of Hydrophobic
;     TITLE OF INVENTION:  Polypeptides
;     NUMBER OF SEQUENCES:  3
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Chiron Corporation
;       STREET:  4560 Horton Street
;       CITY:  Emeryville
;       STATE:  California
;       COUNTRY:  U.S.A.
;       ZIP:  94608
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/477,410A
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Chang, Ling-Fong
;       REGISTRATION NUMBER:  36,482
;       REFERENCE/DOCKET NUMBER:  960.001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (510) 923-2704
;       TELEFAX:  (510) 655-3542
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  166 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US 08 477 410A 1


Alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:        0
 Percent Similarity:  81.818   Percent Identity: 77.273

Alignment_block:
 US-09-832-658-7 x US-08-477-410A-1    ..

 Align seq 1/1  to: US 08 477 410A 1  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||::|||      ||||||  ||||||  ||||||||||||||||||
        1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

Seq_name: /usr/u5/pl/Data/Output/A_7Mb.pep 'S' 1-422 vi.A 1

Seq_documentation_block:
; Sequence 1, Application US/08912768
; Patent No. 6127332
;  GENERAL INFORMATION:
;     APPLICANT:  Goelz, Susan E
;     APPLICANT:  Cate, Richard L
;     APPLICANT:  Pepinsky, Blake R
;     APPLICANT:  Chow, Pingchang E
;     TITLE OF INVENTION:  No. 6127332el Muteins of IFN-Beta
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  James F. Haley, Jr.
;       STREET:  Fish & Neave, 1251 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10020-1104
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/912,768
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/475,774
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haley Jr., James F
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  B179
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 596-9000
;      TELEFAX:  (212) 596-9090
;    INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-912-768-1


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:       0
 Percent Similarity:  81.818   Percent Identity:  77.273

alignment_block:
 US-09-832-658-7 x US-08-912-768-1    ..

  Align Seq 1/1  to: US-08-912-768-1  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||:::|||      ||||||  ||||||  ||||||||||||||||||
        1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:US-09-397-992A-7

seq_documentation_block:
; Sequence 7, Application US/09397992A
; Patent No. 6329175
; GENERAL INFORMATION:
;  APPLICANT: Conklin, Darrell
;  APPLICANT:  Grant, Francis J.
;  APPLICANT:  Rixon, Mark W.
;  APPLICANT:  Kindsvogel, Wayne
;  TITLE OF INVENTION: Interferon-epsilon
;  FILE REFERENCE: 98-46
;  CURRENT APPLICATION NUMBER: US/09/397,992A
;  CURRENT FILING DATE:  1999-09-16
;  PRIOR APPLICATION NUMBER: 60/101,012
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/118,578
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/142,756
;  PRIOR FILING DATE: 1999-07-08
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-397-992A-7


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:  4.944                Gaps:       0
 Percent Similarity:  81.818   Percent Identity:  77.273

alignment_block:
 US-09-832-658-7 x US-09-397-992A-7    ..

  Align Seq 1/1  to: US-09-397-992A-7  from: 1  to: 166

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||:::|||      ||||||  ||||||  ||||||||||||||||||
        1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       17 sGlnLysLeuLeuTrp 22

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US95-03206-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9503206
;  GENERAL INFORMATION:
;    APPLICANT:  Biogen, Inc.
;    APPLICANT:  Goelz, Susan E
;    APPLICANT:  Cate, Richard L
;    APPLICANT:  Pepinsky, Blake R
;    APPLICANT:  Chow, Pingchang E
;    TITLE OF INVENTION:  Novel Muteins Of IFN-Beta
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James F. Haley, Jr.
;      STREET:  Fish & Neave, 1251 Avenue of the
;      STREET:  Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10020-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03206
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haley Jr., James F
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  B179
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 596-9000
;      TELEFAX:  (212) 596-9090
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
PCT US95-03206-1


alignment_scores:
            Quality:   89.00              Length:      22
```

```
                Ratio:  4.944                 Gaps:      0
 Percent Similarity:  81.818   Percent Identity: 77.273

Alignment Block:
US-09-832-658-7 x [illegible]

Align seq 1/1  to:  [illegible]

      21 ATGCCTTACGCCCCTCTTCGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
         ||:::|||      ||||||  ||||||  |||||||||||||||||
       1 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 17

      71 TCAGAAGCTCCTGTGG 86
         ||||||||||||||||
      38 sGlnLysLeuLeuTrp 43

[illegible]

seq_documentation_block:
[illegible]

aligned_scores:
            Quality:   89.00              Length:     22
              Ratio:  4.944                 Gaps:      0
 Percent Similarity:  81.818   Percent Identity: 77.273

alignment_block:
 US-09-832-658-7 x US-08-026-758-22   ..

 Align seq 1/1  to: US-08-026-758-22  from: 1  to: 187

      21 ATGCCTTACGCCCCTCTTCGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
         ||:::|||      ||||||  ||||||  |||||||||||||||||
      22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 38

      71 TCAGAAGCTCCTGTGG 86
         ||||||||||||||||
      38 sGlnLysLeuLeuTrp 43

seq_number: [illegible]

seq_documentation_block:
; Sequence 4, Application US/08912768
; Patent No. 6127332
;  GENERAL INFORMATION:
;    APPLICANT:  Goelz, Susan E
;    APPLICANT:  Cate, Richard L
;    APPLICANT:  Pepinsky, Blake R
;    APPLICANT:  Chow, Pinchang E
;    TITLE OF INVENTION:  No. 6127332el Materns of IFN-Beta
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James F. Haley, Jr.
;      STREET:  Fish & Neave, 1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10020-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/926,758
;      FILING DATE:  19980305
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 578002lman F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126 096 0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413 2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID No:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  187 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  22..187
;      OTHER INFORMATION:  /note  "Hu-IFN-beta"
US-08-026-758-22


aligned_scores:
            Quality:   89.00              Length:     22
              Ratio:  4.944                 Gaps:      0
 Percent Similarity:  81.818   Percent Identity: 77.273

alignment_block:
 US-09-832-658-7 x US-08-026-758-22   ..

 Align seq 1/1  to: US-08-026-758-22  from: 1  to: 187

      21 ATGCCTTACGCCCCTCTTCGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
         ||:::|||      ||||||  ||||||  |||||||||||||||||
      22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 38

      71 TCAGAAGCTCCTGTGG 86
         ||||||||||||||||
      38 sGlnLysLeuLeuTrp 43

seq_number: [illegible]

seq_documentation_block:
; Sequence 4, Application US/08912768
; Patent No. 6127332
;  GENERAL INFORMATION:
;    APPLICANT:  Goelz, Susan E
;    APPLICANT:  Cate, Richard L
;    APPLICANT:  Pepinsky, Blake R
;    APPLICANT:  Chow, Pinchang E
;    TITLE OF INVENTION:  No. 6127332el Materns of IFN-Beta
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James F. Haley, Jr.
;      STREET:  Fish & Neave, 1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10020-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/912,768
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/475,774
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bailey Jr., James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-768-3

alignment_scores:
Quality: 89.00 Length: 22
Ratio: 4.944 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x US-08-912-768-3 ..

Align_Seq 1/1 to: US-08-912-768-3 from: 1 to: 187

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
||||:|||| |||||| |||||| ||||||||||||||||||
22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnC, 38

71 TCAGAAGCTCCTGTGG 86
||||||||||||||||
38 sGlnLysLeuLeuTrp 43

seq_name: /cgn2_6/ptodata/2/iaa/6R_COMB_pep_US-08-406-030A-30

seq_documentation_block:
; Sequence 9, Application US/09206903A
; Patent No. 6200780
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Godowski, Paul J.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Dong-Xiao
; TITLE OF INVENTION: NOVEL TYPE I INTERFERONS
; FILE REFERENCE: P1224-2R1
; CURRENT APPLICATION NUMBER: US/09/206,903A
; CURRENT FILING DATE: 1998-12-07
; PRIOR APPLICATION NUMBER: US 60/106,463
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 9
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-206-903A-9

alignment_scores:
Quality: 89.00 Length: 22
Ratio: 4.944 Gaps: 0
Percent Similarity: 81.818 Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x US-09-206-903A-9 ..

Align_Seq 1/1 to: US-09-206-903A-9 from: 1 to: 187

21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
||||:|||| |||||| |||||| ||||||||||||||||||
22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnC, 38

71 TCAGAAGCTCCTGTGG 86
||||||||||||||||
38 sGlnLysLeuLeuTrp 43

seq_name: /cgn2_6/ptodata/2/iaa/6R_COMB_pep_US-08-406-030A-30

seq_documentation_block:
; Sequence 30, Application US/08406030A
; Patent No. 6270989
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Hauqe, Brian M.
; APPLICANT: Selden, Richard F
; TITLE OF INVENTION: Protein Production and Delivery
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,030A
; FILING DATE: 17-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,391
; FILING DATE: 13-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,586
; FILING DATE: 03-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,583
; FILING DATE: 10-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,840
; FILING DATE: 05-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,188
; FILING DATE: 05-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11704
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09627
; FILING DATE: 05-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: TKT95-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-406-030A-30

```
alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:   4.944                Gaps:       0
 Percent Similarity:   81.818   Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x US-09-206-935-7    ..

 Align seq 1/1  to: US-09-206-935-7  from: 1  to: 187

       21 ATGGCTTACGCCGCTCTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||::|||      ||||||  ||||||  ||||||||||||||||||
       22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 68

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       69 sGlnLysLeuLeuTrp 83

Seq_number: 7 Seq_type: DNA Seq_length: 187 Seq_pep: US-09-206-935-7

Seq_Documentation_block:
; Sequence 7, Application US/09202122
; Patent No. 6210804
; GENERAL INFORMATION:
;  APPLICANT: Chen, Jian
;  APPLICANT: Sodowski, Paul
;  APPLICANT: Wood, William I.
;  APPLICANT: Zhang, Dona Xiao
;  TITLE OF INVENTION: HUMAN INTERFERON EPSILON, A TYPE I INTERFERON
;  FILE REFERENCE: P1224R2 (filed)
;  CURRENT APPLICATION NUMBER: US/09/206,935
;  CURRENT FILING DATE: 1998-12-07
;  EARLIER APPLICATION NUMBER: 60/084,045
;  EARLIER FILING DATE: 1998-05-04
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 187
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-206-935-7


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:   4.944                Gaps:       0
 Percent Similarity:   81.818   Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x US-09-206-935-7    ..

 Align seq 1/1  to: US-09-206-935-7  from: 1  to: 187

       21 ATGGCTTACGCCGCTCTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||::|||      ||||||  ||||||  ||||||||||||||||||
       22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 68

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       69 sGlnLysLeuLeuTrp 83

Seq_number: 7 Seq_type: DNA Seq_length: 187 Seq_pep: US-09-206-936-7

Seq_Documentation_block:
; Sequence 7, Application US/09206936A
; Patent No. 6300475
; GENERAL INFORMATION:
;  APPLICANT: Chen, Jian
;  APPLICANT: Wood, William I.
;  TITLE OF INVENTION: No. 6300475e1 Interferon
;  FILE REFERENCE: P1224R1
;  CURRENT APPLICATION NUMBER: US/09/206,936A
;  CURRENT FILING DATE: 1998-12-07
;  EARLIER APPLICATION NUMBER: US 60/067,897
;  EARLIER FILING DATE: 1998-12-08
;  NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 7
;   LENGTH: 187
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-206-936-7


alignment_scores:
            Quality:   89.00              Length:      22
              Ratio:   4.944                Gaps:       0
 Percent Similarity:   81.818   Percent Identity: 77.273

alignment_block:
US-09-832-658-7 x US-09-206-936-7    ..

 Align seq 1/1  to: US-09-206-936-7  from: 1  to: 187

       21 ATGGCTTACGCCGCTCTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          |||::|||      ||||||  ||| ||  ||||||||||||||||||
       22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 68

       71 TCAGAAGCTCCTGTGG 86
          ||||||||||||||||
       69 sGlnLysLeuLeuTrp 83

Seq_number: 7 Seq_type: DNA Seq_length: 187 Seq_pep: US-09-206-937-7

Seq_Documentation_block:
; Sequence 7, Application US/09206935
; Patent No. 6209877
; GENERAL INFORMATION:
;  APPLICANT: Chen, Jian
;  APPLICANT: Sodowski, Paul
;  APPLICANT: Wood, William I.
;  APPLICANT: Zhang, Dona Xiao
;  TITLE OF INVENTION: HUMAN TYPE I INTERFERON
;  FILE REFERENCE: P1224R2US
```

```
Seq_documentation_block:
;Patent No. 5326859
;     APPLICANT: Sudano, Haruo;Muramatsu, Masami;Taniguchi, Tadatsugu
;     TITLE OF INVENTION: DNA AND RECOMBINANT PLASMID
;     NUMBER OF SEQUENCES: 3
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/06/201,459
;       FILING DATE: 27-OCT-1980
;SEQ ID NO:1:
;       LENGTH:187
5326859-1


Alignment scores:
            Quality:    89.00              Length:      22
              Ratio:   4.944                Gaps:       0
 Percent Similarity:  81.818   Percent Identity: 77.273

Alignment_block:
US-09-832-658-7 x 5326859-1    ..

 Align seq 1/1  to: 5326859-1  from: 1  to: 187

       21 ATGGCTTACGCCGCTCTTGGAGCCCTACAAGCTTCTAGCAATTTTCAGTG 70
          ||:::|||     |||:||   ||||||   ||||||||||||||||||
       22 MetSerTyrAsnLeuLeuGlyPheLeuGlnArgSerSerAsnPheGlnCy 38

       71 TCAGAAGCTCCCTGTGG 86
          |||||||||||||||||
       38 sGlnLysLeuLeuTrp 43
```